Information Systems Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Examining the Impact of Television-Program-Induced Emotions on Online Word-of-Mouth Toward Television Advertising

Tingting Nian, Yuheng Hu, Cheng Chen

INFORMATION SYSTEMS RESEARCH
Vol. 32, No. 2, June 2021, pp. 605–632
ISSN 1047-7047 (print), ISSN 1526-5536 (online)
http://pubsonline.informs.org/journal/isre

Examining the Impact of Television-Program-Induced Emotions on Online Word-of-Mouth Toward Television Advertising

Tingting Nian,[a] Yuheng Hu,[b] Cheng Chen[c]

[a] Paul Merage School of Business, University of California, Irvine, Irvine, California 92697; [b] College of Business Administration, University of Illinois at Chicago, Chicago, Illinois 60607; [c] Sheldon B. Lubar School of Business, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

Contact: tnian@uci.edu, https://orcid.org/0000-0003-3398-5064 (TN); yuhenghu@uic.edu, https://orcid.org/0000-0003-3665-1238 (YH); chench@uwm.edu, https://orcid.org/0000-0002-2884-4171 (CC)

Received: September 5, 2017
Revised: January 11, 2019; February 28, 2020; September 7, 2020
Accepted: September 30, 2020
Published Online in Articles in Advance: February 17, 2021

https://doi.org/10.1287/isre.2020.0985

Abstract. The proliferation of social media platforms allows marketers to gauge consumers' opinions toward brands directly from online word of mouth (WOM). In this paper, we exploit a large-scale television program (the Super Bowl 2016) and exogenous game events that may cause a positive emotional context for fans of the winning team while creating a negative emotional context for fans of the losing team to investigate the impact of television-program-induced emotions on viewers' online WOM behavior toward ads that were aired during program breaks. The results obtained from a difference-in-differences analysis with coarsened exact matching generally support our hypotheses on the direct and congruence effects of television-program-induced emotions. Findings on the direct effect suggest that television-program-induced emotional shocks have a significant effect on the arousal and valence of viewers' online WOM toward ads. We additionally find that a match between television-program-induced emotional shocks and the emotional content of ads leads to a more significant increase in the arousal and more favorable valence of online WOM responses to ads. We discuss the implications of our findings for advertisement design and media planning strategies.

History: Yong Tan, Senior Editor; Gal Oestreicher-Singer, Associate Editor.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/isre.2020.0985.

Keywords: social media • word-of-mouth • television advertising • cross-media effects • social TV

1. Introduction

Social media channels such as Twitter have emerged as the de facto platforms for information sharing and communication, particularly during many television programs and events such as The Voice and the Super Bowl. Such joint consumption of television programs and the production of real-time social media responses is also known as social television (TV), which has become increasingly popular over the past few years. According to a recent report, 72% of Twitter users tend to post tweets about television programs while watching them live (Midha 2014).

Social TV provides great opportunities to analyze the online word of mouth (oWOM) toward television ads. By doing so, advertisers and marketers can obtain important feedback about the strengths and weaknesses of their ads, assess viewers' attitudes toward the advertised brand, and influence viewers' purchase decisions (Wang et al. 2002, Tsang et al. 2004, Hinz et al. 2016). Despite the apparent practical importance of social TV for television advertising, research in this area is still in its infancy. Previous research has shown initial links between television programs and social media conversation (Benton and Hill 2012, Nielsen 2015) and between television advertising and oWOM (Fossen and Schweidel 2016). However, to date, little research has studied the interplay between television programs, ads, and oWOM toward ads.

In this work, we aim to examine and understand this joint relationship, presenting a comprehensive investigation into the social TV activity. In particular, we focus on television-program-induced emotional context for oWOM and address two research questions: First, a television program may evoke emotions from its viewers. For example, a comedy makes viewers laugh and a tragedy makes them cry. Because television ads are aired during the program's breaks, *to what extent do television-program-induced emotions impact the oWOM about ads in terms of arousal and emotional valence?* Moreover, because an ad can also evoke emotions from its viewers, *to what extent do television program–induced emotions interact with ad-induced emotions to jointly impact the oWOM about ads?*

We answer these research questions through the lens of Twitter posts in response to ads aired during a Super Bowl game. The Super Bowl is arguably one of the most-watched annual sporting events and one of the largest television advertising events in the world with an average 110 million viewership and worthy billions each year.[1] During its broadcasting, viewers post actively on social media platforms including Twitter and Facebook about what they saw on television, such as game events (e.g., "Wow, another touchdown by Broncos") and ads (e.g., "that Pepsi ad is not funny at all!"). These viewers, based on their stances and perspectives, tend to have different emotional reactions to the game. For example, a team's touchdown is certainly good news to its fans, whereas it frustrates the other team's fans. On the other hand, nonfans probably would be less emotionally attached to the game because they may not care too much about the outcome of the game. Therefore, we first propose that after the airing of an ad, emotion shocks induced by a subsequent Super Bowl game event (e.g., a touchdown) can carry over, more directly affecting the oWOM communication toward that ad for the fans than for the nonfans. Moreover, ads may induce emotions from their viewers because they are designed to appeal to the viewers' emotions. Thus, we further posit a congruence effect such that a match between television-program-induced emotions and ad-induced emotions will lead to an intensified effect on the online WOM communication toward ads.

To examine our propositions, we construct a multisource data set that includes all 52 ads aired during the Super Bowl 2016 between the Denver Broncos and the Carolina Panthers and more than 1.2 million tweets posted by viewers mentioning these 52 ads and their associated brands. We also measure the arousal and valence of these tweets based on tweet volume and through text analysis. To identify the causal effect of television-program-induced emotional context on viewers' online WOM communication toward ads in terms of arousal and emotional valence, we first identify viewers in our data set who follower either team on Twitter (we call them fans) as the treatment group, because their emotions are likely to be affected by the events and outcomes of the Super Bowl. On the other hand, the control group consists of nonfans (viewers in our data set who do not follow either team on Twitter). Next, we adopt a difference-in-differences (DID) approach that compares changes in the online WOM of the treatment group and control group toward the ad after an emotion-inducing game event (e.g., a touchdown, an interception, or a fumble).

Our analyses identify several interesting results: First, we find a television-program-induced emotion shock, be it positive or negative (with respect to the fan's stance), significantly increases the arousal of oWOM communication about the ad. These emotion shocks also affect the emotional valence of viewers' WOM toward the ad. Such effects are more prominent in the case of negative emotion shocks wherein the treatment group perceives ads less favorably than the control group after being exposed to negative emotion shocks. Interestingly, we further find that a match between television-program-induced emotions and ad-induced emotions significantly intensifies the effect on the arousal of oWOM and also leads to more favorability in online WOM toward the ad.

2. Background

2.1. Social TV and Cross-Media Effects

Our work is inspired by the information systems and marketing literature on cross-media effects. Cross-media effects often refer to the impact of the distribution of content or messages (e.g., music, text, pictures, ads) across different media channels (Forman et al. 2009, Zentner et al. 2013). Among past studies, the most relevant to us are the ones that found television programs or ads influence online searches and shopping behaviors (Lewis and Reiley 2014, Hinz et al. 2016). More recently, a growing body of literature has begun to examine social TV, a cross-media activity in which viewers share their thoughts with other viewers on social media platforms while watching live television. Past work presented initial evidence that linked oWOM with a television program's content. For example, Benton and Hill (2012) showed that when messages are posted on the television screen during the show, they are much more likely to be discussed by viewers on Twitter. More recently, researchers started to explore the relationship between oWOM and television advertising. For example, Hill et al. (2012) analyzed oWOM during the 2012 Super Bowl and found that the level of online consumer engagement, measured by new followers attracted on Twitter during the Super Bowl, can be linked to whether the advertised brand had a social media strategy or not. Chang (2019) examined viewers' emotional reactions on social media during Super Bowl 50 and found that viewers were positive when their team scored but expressed negative emotions when the opposite team scored. Moreover, when a team scored soon after the opposite team scored, viewers expressed a surge of positive or negative emotions, accordingly.[3] Fossen and Schweidel (2016) further showed that some ad characteristics, such as whether a celebrity is in the ad, can have an impact on the oWOM about the advertised brand.

Despite this progress, to date, little research has investigated viewers' social TV activity toward ads in the context of television viewing. In reality, television programs and advertising are naturally connected because ads are usually broadcasted during television programs' commercial breaks. Therefore, our work here aims to extend the current literature on social TV and cross-media effects by exploring the joint relationship between television programs, advertising, and their impact on the volume, arousal, and valence of viewers' oWOM toward ads. In particular, we are interested in how viewers' excitation evoked by the Super Bowl game shocks can be transferred to affect their oWOM toward the ads.

2.2. oWOM

oWOM is a very rich literature. In general, WOM is now widely accepted as crucial to the success of a business (Duan et al. 2008, Mudambi and Schuff 2010). Researchers have placed significant emphasis on the impact of WOM on product demand (Godes and Mayzlin 2004, Chevalier and Mayzlin 2006), firm strategy (Chen and Xie 2008), and market competition (Kwark et al. 2014). Recently, a number of studies have begun to examine the drivers of online WOM. Thus far, researchers have mostly focused on individual factors such as social image, satisfaction, peer effects, as well as social and economic incentives (Forman et al. 2008, Li and Hitt 2008, Toubia and Stephen 2013). Little research, however, has attempted to identify and examine exogenous situational factors that may also influence oWOM behavior. One notable exception is a recent work by Bakhshi et al. (2014), which found the exogenous weather factor can affect Yelp reviews' rating and polarity. On the other hand, research from IS and other fields have extensively studied the role of situational contextual factors such as geolocation and emotion in recommender systems, also known as context-aware recommender systems (Van Setten et al. 2004, Gonzalez et al. 2007, Adomavicius and Tuzhilin 2015). Inspired by these studies, we aim to extend prior literature on oWOM by exploring the impact of an important yet unexplored situational factor—emotions induced by a television program—on viewers' oWOM behavior toward ads aired during the television program.

3. Hypothesis Development

This section presents the main theoretical framework of this paper. Table 1 lists the summary of our hypotheses. We develop our hypotheses under the stimulus (S)–organisms (O)–response (R) framework that is a classic model in the consumer research for capturing elements of the complex process of consumer response (Mehrabian and Russell 1974, Ha and Im 2012, Kim and Johnson 2016). The S-O-R framework demonstrates the effect of external influence (S) on consumers, the internal processes (O) responding to that influence, and the resulting behaviors (R).

Previous research has shown external influences (e.g., advertising, price, product design, or environmental cues) can affect emotional states of the consumers, which further lead to behavioral responses (e.g. intention to act and purchase) (Eroglu et al. 2001, Peng and Kim 2014, Kim and Johnson 2016). The current study develops a research model based on the S-O-R model. Consistent with extant studies, our model considers the television-program-induced emotion shocks and ad-induced emotions as external influence in the stimulus stage. We also consider the viewers' emotional regulation, as the *organism (O)* in our model, which is represented by two dimensions, arousal and pleasure/valence according to the pleasure-arousal (PA) model (Mehrabian and Russell 1974, Russell 1980), and operationalize the response (R) as oWOM communications toward ads.

3.1. Television Program and oWOM Toward Ads

Arousal is the degree to which a person feels excited, stimulated, alert, or active in a situation (Ladhari 2007). A substantial body of research has shown television programs can evoke arousal from viewers through monitoring their heartbeat, respiration rate, and respiration irregularity (Averill 1969, Singh and Churchill 1987, Pavelchak et al. 1988). Relatedly, emotional valence indicates the degree to which a person feels happy, pleased, or contented. Previous research found that an emotionally provoking event can also induce changes in valence simultaneously with the physiological arousal (Myers 2004).

Given its exciting, emotional, and entertaining nature, the Super Bowl no doubt consistently induces arousal and valence from its viewers, affecting their ad appreciation and oWOM toward ads. What is less evident is the direction of the relationship between arousal and valence of viewers' oWOM toward ad and the emotion induced by Super Bowl game shocks, as several theories offer competing perspectives on the phenomenon. Specifically, the excitation transfer theory (Cantor et al. 1975, Isen et al. 1978, Goldberg and Gorn 1987, Tavassoli et al. 1995) suggests residual excitation from a previous stimulus may be carried

Table 1. Summary of Hypotheses

	Direct effect		Congruence effect	
Outcome variable	Positive game shock	Negative game shock	Positive game shock	Negative game shock
Arousal	+ (Hypothesis 1a)	+ (Hypothesis 1a)	+ (Hypothesis 2a)	+ (Hypothesis 2a)
Valence	+ (Hypothesis 1b)	− (Hypothesis 1b)	+ (Hypothesis 2b)	+ (Hypothesis 2b)

over to a subsequent stimulus, amplifying the excitatory response. On the other hand, the cognitive capacity theory states that a positive mood activates information in memory that limits the recipient's processing of other information with less of a positive attitude (Mackie and Worth 1989, Lang et al. 1995). To make sense of these seemingly conflicting perspectives, Norris and Colman (1992) show that ads in print media (e.g., a magazine) can be skipped more easily, as a result of which the appreciation of the emotions induced by print media leads to less ad processing. In other words, the cognitive capacity theory can be applied to explain. However, on radio and television, ads cannot be skipped so easily. As a result, there might be a carryover effect of context appreciation on *forced exposure* ad processing for which the excitation transfer theory can be applied to explain.

Because the Super Bowl ads are broadcasted live on television and often are the focal point of the Super Bowl, we expect that viewers tend not to skip the ads. Therefore, the excitation transfer theory is more suitable. Applying it to our context, after an exogenous game-event emotion shock (e.g., a touchdown or a fumble), we expect viewers' residual excitation can be transferred from the game to intensify their reaction and attitude toward the ads. As a result, these viewers can be very excited about the ads. Meanwhile, for those who are in a positive emotional state they may also evaluate ads more favorably than those in a negative emotional state and vice versa. Finally, we expect the viewers who are fans of the teams in the Super Bowl tend to deeply engage in the game. Thus, they are likely to have a higher arousal and become more emotional than other viewers who are not fans of either team because outcomes of the game events may be of less concern or interest to them. Taken together, we have the following hypotheses.

Hypothesis 1a. *After a positive or negative game-event shock, the arousal of oWOM toward ads posted by the fan groups increases more than that of the nonfan group.*

Hypothesis 1b. *After a positive game-event shock, the emotional valence of oWOM toward ads posted by the fan groups increases more than that of the nonfan group. Conversely, after a negative game-event shock, the emotional valence of oWOM toward ads posted by the fan groups decreases more than that of the nonfan group.*

3.2. Cross-Media Effects on oWOM

Similar to television programs, ads may also impose emotion shocks on viewers because they are designed to appeal to the viewer's emotions (Holbrook and Batra 1987). Because ads are often embedded in television programs or aired during the program timeout/breaks, ad-induced emotions may interact with television-program-induced emotions to jointly influence viewers' WOM toward the ad. These two emotions can be similar or contrasting (e.g., a humorous ad may be placed in a comedy episode program or in a heartbreaking drama episode). The classic mood-congruency theory states that a match between an individual's emotional state and stimuli will lead to a more positive evaluation of the stimuli in the sense that arousal and emotions for the stimuli would be more accessible and heightened than when a mismatch is occurred (Forgas and Bower 1987, Perry et al. 1997, Fiedler et al. 2001, Hinojosa et al. 2009). Thus, we expect that, when an ad induces an emotional response that is congruent with the emotional response induced by a following game event in the Super Bowl (e.g., an individual may feel joyful and positive about the ad, and then the viewer's supported team has a touchdown), the viewer's arousal and emotion toward this ad is intensified.

Whereas the congruent effects of positive emotions are intuitive, the influences of negative emotions, namely, that negative emotional states in viewers should not always result in a negative evaluation of advertising stimuli, are somewhat controversial. One way to explain this is based on the negative-state relief (NSR) model (Carlson and Miller 1987), which suggests people have an innate drive to reduce negative moods. These negative moods can be reduced by engaging in any mood-elevating behavior, including helping behavior and prosocial behavior (Cialdini et al. 1982). Therefore, after watching a dramatic ad that results in a negative mood state for viewers, viewers who later experience a television-program-induced emotional shock that is consistent with their mood state might have a positive reaction to the ad (Kamins et al. 1991).

As discussed earlier, the level of arousal and emotion induced by a game event might be elevated even more for the fan group in the congruent scenario than that in the incongruent scenario. Thus, we have the following hypotheses.

Hypothesis 2a. *If the emotion induced by the game-event shock is congruent with the emotion induced by the ad, the fan groups experience a larger increase in the arousal level of oWOM toward the ad compared with the case when it is not congruent.*

Hypothesis 2b. *If the emotion induced by the game-event shock is congruent with the emotion induced by the ad, the emotional valence of oWOM toward the ad increases more (or decreases less) for fan groups compared with the case when it is not congruent.*

One of the key challenges is how to measure arousal and valence of oWOM. For arousal, Ladhari (2007) provided evidence that arousal has significant influences on consumers' likelihood to generate WOM. Berger (2011), on the other hand, showed arousal drives social transmission of information. In particular, high arousal will boost more sharing than emotions characterized by low arousal. Following these studies, we first use a volume-based method to quantify the impact of arousal on consumers' WOM. Furthermore, arousal is also reflected in the written words as many previous studies have shown (Bradley and Lang 1999, Stevenson et al. 2007, Kuperman et al. 2014). Thus, we also build a text-based measure for the arousal. Although two measures for one variable may be seen as redundant at first glance, they provide different perspectives about the arousal, thus complementing each other to help better understand arousal's effect and the mechanism by which the effect of arousal could manifest (e.g., from the volume perspective, arousal can cause more people to post tweet; from the text perspective, arousal can cause people to use more high-arousal words to express their excitement). On the other hand, we also use a text-based approach to measure emotional valence given that it is one of the most common and popular method in extant emotion analysis (Ding et al. 2008). We provide more details about these measurement in the following section. In sum, our findings are summarized in Table 2.

4. Empirical Methods

4.1. Data Collection

Our multisource data set includes the game progress of the Super Bowl 2016 between Denver Broncos and Carolina Panthers,[4] ads, and Twitter posts about ads and advertised brands. We recorded the airing time of all 52 ads aired during the Super Bowl from 6:30 p.m. to 10:15 p.m. Eastern Standard Time, February 7, 2016. These ads are for 46 brands across eight major product categories such as automobile, food, and beverage industries. During the Super Bowl, ads are aired during mandatory television program commercial breaks and natural breaks caused by on-field events such as field goal attempts, touchdowns, injuries, changes in possession by punts or turnovers, instant replay challenges, and so on.[5] We also obtained the video clips of these ads to measure ad-induced emotions. We supplemented the ad data with oWOM data mentioning these ads. We focus on the oWOM on Twitter because most online conversations about the Super Bowl and its ads occur there. We used Twitter's public Application Programming Interface (API) to capture tweets posted during the Super Bowl that mentioned the official hashtag used for promoting the advertisement (e.g., #MeetTheKetchups). In the online appendix, Table A.1 lists the advertised brands, their official hashtags for the ad (if any), and the volume of tweets about these ads in our data set. For each tweet, we also collected its metadata, including the Twitter user's name, the number of followers, the number of followings, the number of likes, the total number of tweets posted, and which team the user follows, to determine television-program-induced emotions of the user (see the next section). Furthermore, Figure 1 displays changes in the volume of tweets that discussed about those ads aired during the course of the Super Bowl 2016.

4.2. Variable Definitions and Measurements

We define and discuss the key variables in this study. Tables 3 and 4 list these variables used and their summary statistics.

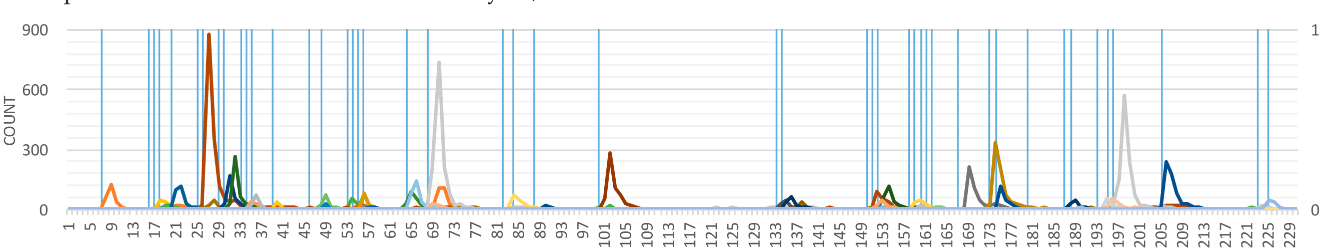
4.2.1. Arousal of Tweets Toward Ads.

We measure the arousal in a tweet which was posted about a specific ad in two different ways.

4.2.1.1. Volume-Based Approach.

First, we use the volume-based approach to measure the arousal of tweets toward ads. This approach is inspired by prior studies that found that high arousal will increase one's sharing activity more than low arousal (Ladhari 2007, Berger 2011). In particular, we count the number of tweets per minute, per ad, and per group (e.g., posted by Broncos' fans, Panthers' fans, and other nonfans). Identifying the stance of viewers can be a challenging task. We follow a simple approach by

Figure 1. (Color online) Volume of Tweets About Each of These 52 ads Aired During the Super Bowl 2016 from 6:30 p.m. to 10:00 p.m. Eastern Standard Time on February 27, 2016



Notes. The vertical lines indicate the airing time for each ad. Ads were aired during the game's commercial or natural breaks, such as field goal tries, touchdown, injury, changes in possession by punts or turnovers, instant replay challenges, and so on. As we can observe from this figure, volume spikes often come right after ads were aired.

Table 2. Summary of Findings

	Direct effect		Congruence effect	
Outcome variable	Positive game shock	Negative game shock	Positive game shock	Negative game shock
Arousal	+ (Hypothesis 1a, yes)	+ (Hypothesis 1a, yes)	+ (Hypothesis 2a, yes)	+ (Hypothesis 2a, yes)
Valence	– (Not significant)	– (Hypothesis 1b, yes)	+ (Hypothesis 2b, yes)	+ (Hypothesis 2b, yes)

using the followers of the Denver Broncos' and the Carolina Panthers' official Twitter accounts (@Broncos and @Panthers) as a proxy for the stance, because following a team's official Twitter account is a strong indicator that a user is a committed fan of that team (Frederick et al. 2012, Gray et al. 2017). To this end, we first collect the full follower lists of the two teams (1.87M for @Broncos and 1.54M for @Panthers at the time of crawling) and check if a Twitter user in our data set follows either team. We find that 85,500 users (9.8%) in our data set follow the Denver Broncos, whereas 107,800 (12.3%) users follow the Carolina Panthers. Aside from the 1.5% of users who follow both teams, which we removed from the study, we classify the rest as nonfans.

4.2.1.2. Text-Based Approach. We also develop a text-based approach using the popular arousal dictionary developed by Warriner et al. (2013) to measure arousal from the text content. This dictionary is widely used in various psycholinguistics analysis and text mining tasks such as Kuperman et al. (2014), Rajadesingan et al. (2015), Mohammad (2016), and Tan et al. (2016). This dictionary contains 13,915 words with arousal scores ranging from 1.6 to 7.79. The lowest score indicates relaxed, calm, sluggish, dull, sleepy, or unaroused, whereas the highest score indicates stimulated, excited, frenzied, jittery, wide awake, or aroused. We use the dictionary to compute arousal scores for each tweet. Specifically, after standard text preprocessing procedures such as stopwords and URL removal, we sum up arousal scores of all the words in a tweet and normalize it with the total number of words in the tweet.

4.2.2. Emotional Valence of Tweets Toward Ads

Because tweets are short and noisy, extracting a tweet's emotional valence or the sentiment is known to be a challenging task. We first manually analyzed a random sample of 200 viewers' tweets to understand the nature of oWOM toward ads in the Super Bowl. We observed that these tweets come in a variety of forms, such as complaints (e.g., *"@Doritos, your ad is senseless #facepalm"*) and compliments (e.g., *"this ad is so hilarious #BaldwinBowl"*). Neutral tweets were rare, probably because viewers tend to express their views about these ads on social media. Therefore, in this study, we consider the emotional valence of these tweets to be binary, for example, either positive or negative.

Table 3. Variable Description

Variable(s)	Description
Log($Volume_{kjt}$)	The logarithm of the number of tweets posted by group k about ad j per minute at time t.
$Arousal_{ikjt}$	The text-based arousal level of a certain tweet i about ad j posted by group k at time t.
$Valence_{ikjt}$	The emotional valence of tweet i about ad j posted by group k at time t.
$GameSentCum_{it}$	The cumulative game-induced emotion scores for a user i at time t.
Log($TimeAfterAd_{jt}$)	The logarithm of the seconds elapsed after the ad's airing until time t.
$GameSentCum_{kt}$	The cumulative game-induced emotion scores for group k at time t.
Log($TimeAfterAd_{jt}$)	The logarithm of seconds elapsed after the ad j's airing until time t.
$CongruentAdSent_j$	1 if the ad-induced emotion is matched with the game-event-induced emotion.
Log($Followers_i$)	The logarithm of the number of followers a user has at the time he posted tweet i.
Log($Followings_i$)	The logarithm of the number of accounts a user follows at the time he posted tweet i.
Log($Likes_i$)	The logarithm of the total likes a user has at the time he posted tweet i.
Log($NumberofTweets_i$)	The logarithm of the total tweets a user has at the time he posted tweet i.
$Treat_i$	1 if the user who posted tweet i supports either Broncos or Panthers.
$Treat_k$	1 if the group k supports either Broncos or Panthers.
$Post_t$	1 if the tweet(s) was/were posted after the positive/negative game shock.

Table 4. Summary Statistics

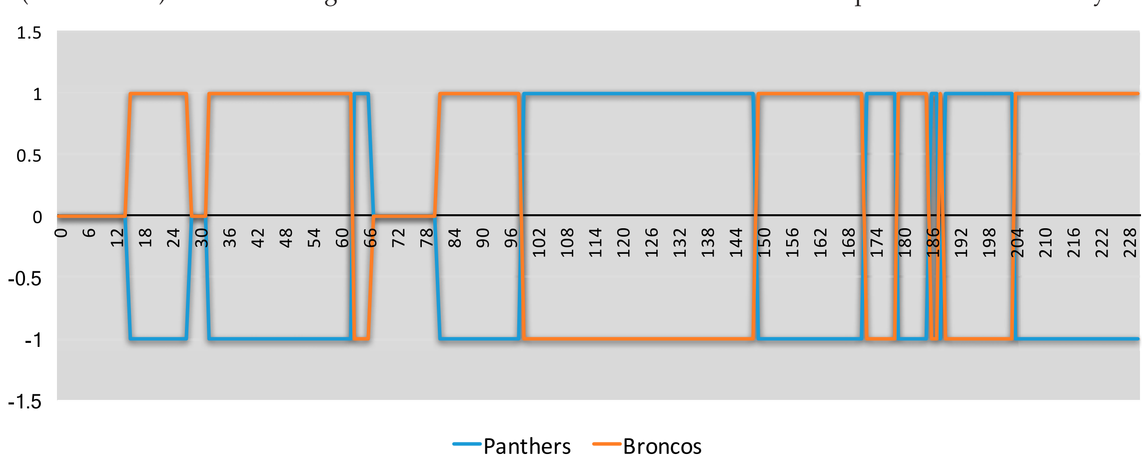
Variable	Mean	Standard deviation	Min	Max	Observations
Volume-based arousal	2.90	19.50	0	1507	53,592
Text-based arousal	1.48	0.98	0	7	53,897
Valence	0.12	0.40	−1	1	53,897
Treat	0.5	0.5	0	1	53,897
Post	0.5	0.5	0	1	53,897
Time after ad (in seconds)	18.57	46.81	0	298	53,897
Cumulative game sentiment	0.00	0.59	−2	2	53,897
Followers (in thousands)	7.785	129.486	0	10.200	53,897
Followings (in thousands)	1.247	6.801	0	518.824	53,897
Likes (in thousands)	6.370	16.069	0	533.132	53,897
Number of tweets (in thousands)	30.829	107.457	0	5,793.257	53,897

To automatically measure the valence of tweets, we develop a Twitter sentiment classifier based on a novel deep learning architecture based on convolutional neural network and long short term memory network, which is able to achieve an overall accuracy of 91.5% in classifying tweets into either positive or negative. We have also tried other classifiers and approaches such as logistic regression, support vector machine (SVM), and a lexicon-based approach and found this deep learning-based approach achieved the most satisfying classification prediction performance (see Online Appendix B for more details). Therefore, we applied our approach to measure the emotional valence for all the tweets in our data set. As a result, we find the average valence score of tweets is 0.12.

4.2.3. Television-Program-Induced Emotions. As discussed earlier, television-program-induced emotions can vary, depending on who is watching the program and the content of the program. For example, when watching the Super Bowl, fans of one team may have a positive emotion after the team has a touchdown, whereas fans of the opposing team may experience a negative emotion. Moreover, one viewer can be in a positive emotion after her team has a touchdown but soon in a negative emotion after her team has a critical turnover. Therefore, to measure the emotional shocks induced by the television program for a viewer, we need to identify (1) the stance of the viewer and (2) the most recent game event that happened before the tweet was posted.[6]

We already identified the stance of the viewer when we measure the tweet volume for different groups (see Section 4.2.1). Next, to measure television-program-induced emotions for these users, we obtain the play-by-play game report from ESPN[7] and the replay of the Super Bowl 2016. We then ask six raters who follow American football games very closely to watch the replay and read the play-by-play game report. Next, we ask them to rate their emotions toward every major event in the game (e.g., touchdown, field goal, fumble, interception) from −1 (negative), 0 (neutral), to +1 (positive) as if they were Broncos' or Panthers' fans. The interrater reliability measured by Fleiss' kappa is $k = 0.96$, indicating a strong agreement among the raters. Finally, we aggregate the

Figure 2. (Color online) Television-Program-Induced Emotions over the Course of the Super Bowl 2016 Rated by Six Raters



Notes. The x axis indicates the game progress. The y axis indicates the aggregated emotions. The Denver Broncos won the Super Bowl against Carolina Panthers with the final score 24-10.

rated emotions by taking the average. Figure 2 shows the aggregated emotions for representing the television-program-induced emotions for fans over the course of the Super Bowl. For example, the Denver Broncos had a touchdown at around the 32nd minute, which our raters rated as a positive shock for Broncos' fans and a negative shock for Panthers' fans. The emotional shocks for both teams' fans flipped when the next major event, a touchdown by the Panthers, occurred. To validate our measure of game-induced emotions, we compare it with the ESPN's in-game win probability: a proxy of a team's chances of winning at any given point in a game, simulated based on the performance of historical teams in the same situation.[8] We found the correlation as high as 0.926. In other words, when Broncos' win probability is boosted after a touchdown, its fans' emotion is also likely to be positive. Thus, our measure for television-program-induced emotions is validated.

4.2.4. Ad-Induced Emotions.

Ads can be very rich in emotional content, yet they often stress one type of emotion over another. For instance, Hyundai's ad "Better" depicts a baby with an exposed V8 engine heart who grows up and works at Hyundai's design studio with a drive to make better cars. Most of the viewers agreed it strikes more of an inspirational note than a humorous tone adopted by many other brands.[9]

We use a two-dimensional model of emotion as a guide to categorize each ad. Although there are other accepted models of emotions, models with more than two dimensions are useful largely for differentiating complex emotions and cognitive states. Instead, we use a two-dimensional model of emotions because we aim to investigate basic emotions from an ad that can match emotions from the television program, and how these emotions can jointly affect consumers' oWOM communication about ads. We categorize each ad into a positive or negative emotional state using the classic positive-activation-negative-activation (PANA) model (Watson et al. 1988). The PANA model has since been used extensively by other researchers in advertising, marketing, management, communication, and so on (Babin et al. 1998, Pugh 2001, Yang and Smith 2009, Elder and Krishna 2011, Botha and Reyneke 2013). In the PANA model, the positive affect (PA) reflects the extent to which a person feels enthusiastic, active, and alert. By contrast, the negative affect (NA) is a general dimension of subjective distress and unpleasant engagement that subsumes a variety of aversive emotional states, including anger, contempt, disgust, guilt, fear, and nervousness.

To measure positive and negative emotions induced by ads, we use six raters to watch and evaluate each ad. After watching an ad, each rater decides which labels would describe his or her feelings and emotions. All raters individually watched the replay of all 52 ads and labeled them using 20 PANA descriptors (e.g., enthusiastic, interested, determined, excited, inspired). The interrater reliability via Fleiss' kappa is $k = 0.89$, indicating the high agreeableness and reliability of the rating results. As a result, if an ad conveys more positive emotions than negative emotions, we treat the ad as a positive ad, and vice versa.

5. Empirical Analysis

In this section, we first present model-free evidence on how viewers' oWOM toward Super Bowl ads is affected by game-induced emotions. Next, we take one step further by estimating the causal effect of these emotion shocks using a difference-in-differences (DID) approach.

5.1. Model-Free Evidence

Table 5 summarizes the average number of followers, followings, tweets, and likes for Broncos' fans, Panthers' fans, and nonfans in our data set. It shows the activity pattern of fans is quite similar to that of nonfans. We use the t test to compare activities of either Broncos' fans or Panthers' fans with those of nonfans. The last two columns show the p values of the t test for each type of activity between Broncos' fans and Panthers' fans, as well as nonfans. Except for the number of tweets, the nonfan group's social media activities are quite similar to those of the fan groups.

To get an overview of the raw treatment effect of television-program-induced emotional shocks, we now present the model-free evidence of changes in viewers' oWOM responses to ads before and after an emotion-inducing game event during the Super Bowl. Results are shown in Figure 3. From Figure 3(a),

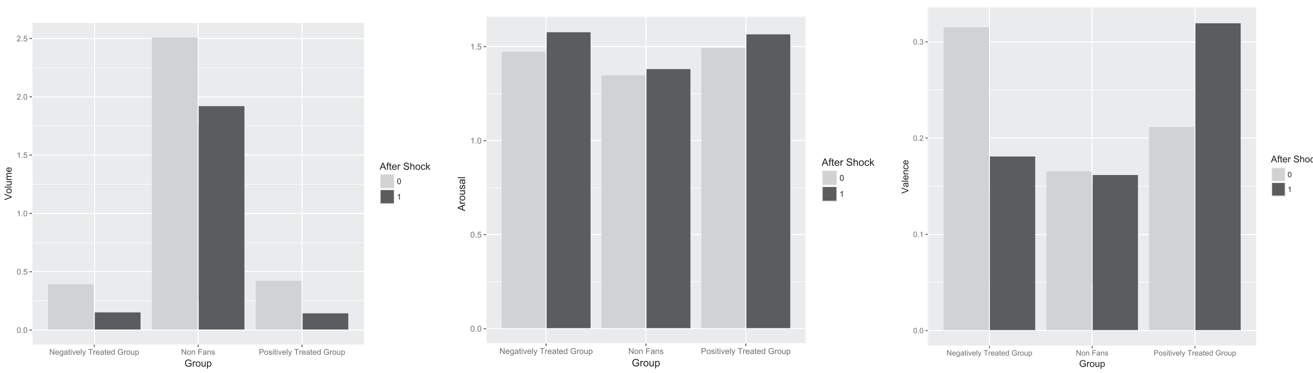
Table 5. Comparing Activity Levels of Fans and Nonfans

Variable	Broncos' fans (B) fans mean	Panthers' fans (P) fans mean	Nonfans (N) follower	t test p value for B-N	t test p value for P-N
Follower	3,650.75	4,264.33	8,618.15	0.367	0.230
Following	1,272.39	1,207.23	1,250.59	0.366	0.831
No. of tweets	16,124.87	20,631.64	32,876.72	0.000***	0.000***
Likes	6,275.77	6,891.068	6,221.574	0.758	0.633

***, **, and * Significance at the 1%, 5%, and 10% levels, respectively.

Figure 3. Before-After Changes in the Volume, Arousal, and Emotional Valence of Tweets by Different Groups of Viewers

(a) (b) (c)



Volume Arousal Valence

we first observe that the average volume of oWOM toward ads is lower after the shock for both fan and nonfan groups. This is expected because of the time wear-out effect, as most ad-related WOM was generated immediately after the ad was aired, and as time elapsed, viewers' interests tended to shift toward other topics such as the game (this effect also resonates well with the sudden spikes and quick drops in the tweets volume over the event timeline shown in Figure 1). Next, from Figure 3(b) we observe that no matter whether television-program-induced emotional shocks were positive or negative, the fan group tended to experience a more heightened level of arousal reflected in their tweet content than the nonfan group. Moreover, in Figure 3(c), by comparing the average valence of their WOM before and after the shock, we observe that the average emotional valence of the oWOM of the leading team's fans was increased after a positive emotional shock. However, for viewers who experienced a negative emotional shock, the emotional valence of their WOM experienced a notable drop. In sum, this model-free evidence is generally consistent with our hypotheses that television-program-induced emotional context has a positive effect on the arousal of viewers' oWOM responses to ads and also affects the valence in a significant way. Next, to control for factors such as time trends, the time wear-out effect and the influence from prior game events and to causally assess the significance of both direct and congruence effects, we turn to the DID analysis.

5.2. DID Analysis

In our DID analysis, we use the exogenous nature of an emotion-inducing event that takes place in the television program. Moreover, we identify viewers who follow either team on Twitter (fans) as the treatment group. As a result, for viewers in this group, their oWOM responses to ads are likely to be affected more by the emotional shock from a game event. On the other hand, the control group consists of nonfans (who do not follow either team). For viewers in this group, their oWOM toward ads are less likely to be affected by the emotional shock. To account for potential systematic differences in the susceptibility to emotion-inducing events between treatment and control groups, we include a dummy for the treatment group, who might be more prone to effects brought by a game-induced emotion shock. We also include a dummy for the postevent period to indicate whether a tweet was posted after a game-induced emotion shock. Although we have data at every second, we divide them into two periods—before and after an emotional shock—to avoid downward bias for the standard errors when using multiperiod datasets (Bertrand et al. 2004).

We first use the following model to measure the causal effect of television-program-induced emotion shocks on the arousal (measured by volume) of viewers' oWOM toward ads:

$$\begin{aligned}
\text{Log}(Volume)_{kjt} &= \beta_0 + \beta_1 Post_t + \beta_2 Treat_k \times Post_t \\
&\quad + \beta_3 Treat_k \times CongruentAdSent_j \\
&\quad + \beta_4 Post_t \times CongruentAdSent_j \\
&\quad + \beta_5 Treat_k \times Post_t \times CongruentAdSent_j \\
&\quad + \beta_6 GameSentCum_{kt} + \beta_7 \text{Log}\big(TimeAfterAd_{jt}\big) \\
&\quad + \alpha_j + \gamma_k + \epsilon_{kjt}.
\end{aligned} \tag{1}$$

Our dependent variable in Equation (1) is the logarithm of the volume of tweets posted by group k per minute, regarding a certain ad j, where the group indicator $k \in$ {Broncos' fans, Panthers' fans, non-fan}.[10] The dummy indicator $Treat_k = 1$ indicates the group is a treated group (e.g., Broncos' or Panthers' fans), and zero if the group is the control group, namely, the nonfan group. The dummy variable $Post_t = 1$ if tweets posted by a group at time t is after the exogenous game-induced emotion shock, and zero if before the game-induced emotion shock. We also capture the time wear-out effect by including $\text{Log}(TimeAfterAd_{jt})$, which denotes the logarithm of the number of seconds elapsed after the ad was aired until time t. In practice, we set a five-minute window after the ad to eliminate other confounding factors that might influence viewers' online WOM responses to the ad.[11] Moreover, each individual event shock may pose some influence on viewers' emotions, and such effects of a sequence of positive and negative events along the course of the game are accumulative. Therefore, we control for the cumulative game-event-induced emotions for group k by including a variable $GameSentCum_{kt}$, which is a cumulative game-event-induced emotion score for users of group k at time t. $CongruentAdSent_j$ equals 1 if the ad-induced emotion is congruent with the emotional shock induced by the game event.[12] We also include ad fixed effects α_j to control for ad-specific effects on viewers' tweets toward ads and group fixed effects γ_k to control for group-specific effects among these teams.[13]

Next, to measure the causal effect of television-program-induced emotion shocks on the text-based arousal and emotional valence of viewers' oWOM, we use the following model depicted in Equation (2). Unlike Equation (1), here we focus on tweet-level analyses because, on average, a user in our sample posted fewer than two tweets.[14]

$$\begin{aligned} y_{ikjt} = {} & \beta_0 + \beta_1 Post_t + \beta_2 Treat_i \times Post_t \\ & + \beta_3 Treat_i \times CongruentAdSent_j \\ & + \beta_4 Post_t \times CongruentAdSent_j \\ & + \beta_5 Treat_i \times Post_t \times CongruentAdSent_j \\ & + \beta_6 GameSentCum_{it} \\ & + \beta_7 \text{Log}(TimeAfterAd_{jt}) + \beta_8 X_i \\ & + \alpha_j + \gamma_k + \epsilon_{ikjt}. \end{aligned} \tag{2}$$

The dependent variables in regression Equation (2) are the text-based arousal level or the emotional valence of the tweet discussing ad j, posted by user i, who belongs to group k at time t. On the right-hand side, $Treat_i$ is an indicator denoting whether user i who posted the tweet belongs to the treatment group. It is set to 1 if the tweet was posted by a user who is a fan of either team and 0 otherwise. We capture the time effect with $Post_t$, an indicator denoting whether the tweet posted at time t was after the treatment period (again we limit the time window to five minutes after the airing of the ad; we also used different lengths of time windows, and found similar results.). It is set to 1 if the tweet was posted after the exogenous game-induced emotion shock, and zero if before the emotion shock. The variables $GameSentCum_{it}$ and $\text{Log}(TimeAfterAd_{jt})$ are defined similarly as in regression Equation (1). The dummy variable $CongruentAdSent_j$ equals 1 if the ad-induced emotion is congruent with the emotional shock induced by the game event. In addition, because this analysis is at the tweet level, we include X_i to control for viewer-specific characteristics such as the number of followers and followings, the number of posts, and the number of likes, which were collected before the game started to minimize the risk of reverse causality. Again, we also include ad fixed effects α_j to control for ad-specific effects on viewers' tweets toward ads and group fixed effects γ_k to control for group-specific effects among these teams.

Our key variables of interest are the interaction terms $Treat_i \times Post_t$ and $Treat_i \times Post_t \times CongruentAdSent_j$. As long as no other systematic reason exists for why the group of fans should change their emotional responses to ads differently from the group of nonfans after the game-induced emotion shock, we can interpret the interaction coefficient β_2 as measuring the causal impact of the television-program-induced emotion shock on viewers' oWOM communication toward ads. Given the exogenous nature of the emotional content of the ad (Super Bowl ads are typically designed and planned months ahead of time by brands themselves), we can also interpret the coefficient β_5 as the congruence effect of ad-induced emotions and matched television-program-induced emotions on viewers' oWOM responses to ads.

5.3. Results

We begin by reporting the main effects of television-program-induced emotion shocks on the volume-based arousal level of viewers' oWOM responses to an ad in Table 6. We separately estimate the impact of positive and negative television-program-induced emotion shocks, because they might affect viewers differently (Berger et al. 2010). In both cases of positive emotion shocks and negative emotion shocks (columns (1)–(4)), coefficients on both $Post_t$ and $\text{Log}(TimeAfterAd_{jt})$ are negative, consistent with our observation that tweets were posted immediately after the ad was aired, and the time wear-out effect is significant here. One of our independent variables of key interest is the interaction term $Treat_k \times Post_t$, which shows statistically significant and positive coefficients in columns (1) and (3). This finding suggests that after television-program-induced emotion shocks, no matter they are positive or negative, the treated group experienced a greater boost in the arousal level compared with the control group, evidenced by the greater increase in tweets generated. This result provides strong support for Hypothesis 1a. However, we observe that the coefficient on the three-way interaction term among $Treat_k$, $Post_t$ and matched ad-induced emotions is statistically significant and

Table 6. Effects of Television-Program-Induced Emotional Shocks on the Volume of Tweets Toward Ads

	Positive game shock		Negative game shock	
	(1)	(2)	(3)	(4)
Post	−0.771***	−0.690***	−0.774***	−0.786***
	(0.036)	(0.078)	(0.036)	(0.037)
Treat×Post	1.041***	0.829***	1.058***	1.092***
	(0.038)	(0.099)	(0.038)	(0.041)
Treat×Congruent Ad Sentiment		−0.219**		−0.217**
		(0.099)		(0.099)
Post×Congruent Ad Sentiment		−0.094		−0.082
		(0.081)		(0.081)
Treat×Post×Congruent Ad Sentiment		0.247**		−0.237**
		(0.107)		(0.107)
Cumulative Game Sentiment	0.224***	0.220***	−0.107***	−0.103***
	(0.038)	(0.038)	(0.038)	(0.038)
Log(time after ad)	−0.438***	−0.438***	−0.437***	−0.437***
	(0.010)	(0.010)	(0.010)	(0.010)
Ad fixed effects	Yes	Yes	Yes	Yes
Group fixed effect	Yes	Yes	Yes	Yes
Adjusted R^2	0.68	0.68	0.68	0.68
Observations	15,608	15,608	15,608	15,608

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The dependent variable is the logarithm of the volume of tweets posted by a certain group of users per minute. *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is not a fan of any NFL teams including Broncos or Panthers. The *Post* takes the value of one if the tweets were posted after a major game event. The *Congruent Ad Sentiment* is an indicator that takes a value of one if the game-induced emotion is congruent with the ad-induced emotion and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

positive in column (2) and is significantly negative in column (4). Although the treatment group responded to positive and negative game-induced shocks in the same way, a more heightened level of arousal and more tweets generated than the control group, it responded differently to the congruence effect. When game-induced emotions matched with ad-induced emotions, the treatment group experienced an even more intensified level of arousal because of the congruence effect in the case of positive game-induced emotion shocks. However, the opposite is true in the case of negative game-induced emotion shocks. Therefore, using volume as a measure of arousal, we only find supportive evidence to our Hypothesis 2a in the case of positive shocks.

Next, we present results on the effects of game-induced emotion shocks on the text-based arousal of viewers' oWOM toward an ad in Table 7. In general, columns (1) and (3) show that after a game event that induced a positive/negative emotion shock, both the treated and control groups witnessed an increased level of arousal. Furthermore, the fan group experienced a larger increase in the arousal level of their oWOM responses to ads compared with the nonfan group. More interestingly, when the game-induced emotions matched with the ad-induced emotions, the fan group experienced an even greater increase in the arousal level compared with the case when these two were not matched, as shown in the coefficients on $Treat_i \times Post_t \times CongruentAdSent_j$ in columns (2) and (4). These results provide strong support for Hypotheses 1a and 2a.

Finally, we examine the effect of television-program-induced emotion shocks on the emotional valence of viewers' WOM toward the ad. Table 8 displays the results, which provide consistent and significant evidence that, after a negative program-induced emotion shock, viewers of the treated team experience an even larger decline in the emotional valence of their online WOM toward ads compared with viewers of the control group, as shown in column (3). However, $Treat_i \times Post_t$ is not significant in the case of positive emotion shocks. Regarding the congruence effect, coefficients on the interaction term, $Treat_i \times Post_t \times CongruentAdSentiment_j$, in both cases of positive and negative emotional shocks, are statistically significant and positive. This finding indicates that when game-induced emotions matched with ad-induced emotions, viewers tended to view the ad more favorably, which strongly supports our Hypothesis 2b. Again, the coefficient on Log($TimeAfterAd_{jt}$) reveals a wear-out

Table 7. Effects of Television-Program-Induced Emotional Shocks on the Arousal of Tweets Toward Ads

	Positive game shock		Negative game shock	
Post	0.025*	–0.054***	0.032**	–0.076***
	(0.015)	(0.020)	(0.015)	(0.020)
Treat × *Post*	0.147***	0.027	0.171***	0.003
	(0.023)	(0.032)	(0.022)	(0.035)
Treat × *Congruent Ad Sentiment*		0.021		–0.249***
		(0.024)		(0.044)
Post × *Congruent Ad Sentiment*		0.096***		0.125***
		(0.016)		(0.017)
Treat × *Post* × *Congruent Ad Sentiment*		0.160***		0.212***
		(0.032)		(0.034)
Cumulative Game Sentiment	–0.053*	–0.056**	0.072***	0.052*
	(0.028)	(0.028)	(0.027)	(0.027)
Log(time after ad)	–0.024***	–0.018***	–0.029***	–0.023***
	(0.005)	(0.005)	(0.005)	(0.005)
Log(followers)	–0.011**	–0.010**	–0.016***	–0.016***
	(0.004)	(0.004)	(0.004)	(0.004)
Log(followings)	–0.000	–0.001	0.003	0.004
	(0.005)	(0.005)	(0.005)	(0.005)
Log(likes)	–0.001	–0.000	–0.001	–0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Log(number of tweets)	–0.011***	–0.012***	–0.010***	–0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Ad fixed effects	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.05	0.05	0.05	0.05
Observations	41,930	41,930	43,072	43,072

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The dependent variable is the arousal score of a tweet posted by a certain group of users. *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is not a fan of Broncos or Panthers, and is not a fan of any of other NFL teams. The *Post* takes the value of one if the tweet was posted after a positive or negative game shock. The *Congruent Ad Sentiment* is an indicator which takes a value of one if the game-induced emotion is congruent with the ad-induced emotion, and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.
***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

effect over time: as the time after the airing of an ad elapsed, viewers' response to the ad became less positive. The negative coefficient on $Post_t$ might be because TV events may be distracting (Benton and Hill 2012).[15]

The variation inflation factor (VIF) of these regressions are no larger than three, so multicollinearity is not an issue. Taken in sum, these estimates provide strong evidence supporting Hypotheses 1a, 2a, and 2b. Hypothesis 1b is only supported in the case of negative emotion shocks. The effects of television-program-induced shocks differ in the positive versus negative emotion shock scenarios, and viewers seem to respond more to negative emotion shocks than to positive emotion shocks. This makes sense and corresponds well to the so-called *negativity bias*. In other words, things of a more negative nature (e.g., unpleasant thoughts, emotions; harmful/traumatic events) have a greater effect on one's psychological state and processes than neutral or positive things (Baumeister et al. 2001, Rozin and Royzman 2001).

6. Robustness Checks

In this section, we conduct a series of additional analyses to test the robustness of our main findings. We begin by providing convergent evidence for our main findings by constructing a different control group, using matching methods, and measuring valence in a different way. Next, we examine the robustness of our results by using falsification tests such as pretrend analyses and placebo tests. Finally, we repeat our analyses on a newly collected data set of ads and tweets during the Super Bowl 2017.

6.1. Other Teams' Fans as a Control Group

To estimate the causal effect of television-program-induced emotion shocks on viewers' oWOM behavior toward the ad, in our previous analysis, we assume

Table 8. Effects of Television-Program-Induced Emotional Shocks on the Valence of Tweets Toward Ads

	Positive emotional shock		Negative emotional shock	
	(1)	(2)	(3)	(4)
Post	−0.019***	0.035***	−0.017***	0.032***
	(0.006)	(0.008)	(0.006)	(0.008)
Treat × *Post*	−0.008	−0.063***	−0.034***	−0.013
	(0.010)	(0.024)	(0.010)	(0.011)
Treat × *Congruent Ad Sentiment*		−0.015		−0.054***
		(0.011)		(0.018)
Post × *Congruent Ad Sentiment*		−0.067***		−0.059***
		(0.007)		(0.007)
Treat × *Post* × *Congruent Ad Sentiment*		0.066**		0.127***
		(0.026)		(0.028)
Cumulative Game Sentiment	0.002	0.003	0.019***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)
Log(time after ad)	−0.028***	−0.031***	−0.030***	−0.033***
	(0.002)	(0.002)	(0.002)	(0.002)
Log(followers)	0.000	0.000	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Log(followings)	0.000	0.000	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Log(likes)	0.005***	0.004***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Log(number of tweets)	−0.009***	−0.008***	−0.007***	−0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Ad fixed effects	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes
Adjusted r^2	0.06	0.07	0.06	0.07
Observations	41,930	41,930	43,072	43,072

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The group of control users are those who do not support any NFL team. The dependent variable is the emotional valence of a tweet posted by a certain user. The variable *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is not a fan of any NFL teams including Broncos or Panthers. *Post* takes the value of one if the tweet was posted after the exogenous emotional shock, such as a touchdown. The *Congruent Ad Sentiment* is an indicator which takes a value of one if the game-induced emotion is congruent with the ad-induced emotion, and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

the nonfan group could serve as a valid counterfactual for the fan group. However, the fan group and nonfan group might differ in systematic ways in their oWOM behavior. Therefore, we use fans of the other 30 NFL teams that did not make it to the Super Bowl as an alternative control group.[16] Using fans of other teams could help reduce the selection bias to some extent. We replicate our analysis using tweets posted by this group of users. Results are displayed in Tables 9 and 10.

Coefficient estimates on $Treat_i \times Post_t$ in columns (1) and (2) of Tables 9 and 10, again, are statistically significant and positive, which implies that after a positive/negative game-induced emotional shock, the fan groups are more incentivized to generate online WOM than the nonfan groups. Moreover, the coefficient estimate of the three-way interaction term is statistically significant and positive in both cases of positive and negative game-induced emotional shocks. In columns (3) and (4) of Tables 9 and 10, we observe that the fan group indeed experienced a boost in their arousal level, compared with the nonfan groups, and the boost is even bigger when game-induced emotions matched with ad-induced emotions. Therefore, with the alternative control group, we found support for Hypothesis 1a and Hypothesis 2a using both the volume-based and the content-based measures of arousal. Most of the results in columns (1)–(4) are consistent with earlier findings, whereas the coefficient on $Treat_i \times Post_t$ loses its significance in column (5) of both tables: a result possibly because of the much smaller sample size. However, we still found statistically significant and positive coefficients on the three-way interaction term in column (6), which supports Hypothesis 2b.

To ensure that fans of the other 30 NFL teams are indeed *neutral* to serve as a valid control group, we eliminate users who might be against one of these two teams (Broncos and Panthers), although they are fans of the other 30 NFL teams and do not follow either of these two teams. Specifically, we performed two layers of filters. First, we removed followers of any of the top three rivalry teams of Broncos and Panthers, which are ranked based on rivalry scores from Tyler and Cobbs (2017). On top of that, we further eliminated Twitter users with historical Tweets mentioning any of the aforementioned rivalry teams. Results are displayed in Table A.7 in the online appendix, which are similar to those obtained in Tables 9 and 10. We also constructed samples by removing users who follow the top one or two rivalry teams of the competing teams. The results are similar.

6.2. Coarsened Exact Matching

Given the descriptive statistics in Table 5, systematic differences may exist in social media activities among Broncos' fans, Panthers' fans, and the nonfan group. Therefore, we apply a coarsened exact match (CEM) (Iacus et al. 2011) method to address concerns related to heterogeneity between the treatment group and control group, and to control for the confounding influence of pretreatment control variables that might affect our outcome of interest. In contrast to propensity score matching, which reduces a multidimensional vector into a single dimension, CEM allows researchers to independently and exactly match groups across multiple attributes. CEM works by first *coarsening* all the observations into strata based on covariates and retaining only strata that contain both treated and control observations (Aggarwal and Hsu 2013). By doing so, it automatically restricts the matched data to areas of common empirical support.

We apply the CEM routine and match on viewers' demographics based on four covariates: the number of followers, the number of followings, the number of likes, and the number of tweets posted, as well as the ad that each tweet mentions. The multivariable imbalance statistic (multivariate $L1$ distance) is reduced from 0.326 to 0.279, using the CEM matching procedure. The variable-to-variable comparison on $L1$ distance is shown in Table A.8 in the online appendix, which suggests the imbalance measure of most variables improved after CEM matching. We fit the regression model in Equations (1) and (2) using the matched sample. Results shown in Table 11 provide similar estimates and support for our previous findings. The game-event-induced emotion shocks significantly increased the arousal level in treated viewers' oWOM toward the ad in both cases of positive emotion shocks and negative emotion shocks. When game-event-induced emotion shocks matched with the emotional tone of ads, viewers tended to respond more favorably to ads, and also experienced a greater increase in the arousal level in both cases of positive and negative emotion shocks. Although supportive evidence is lacking for the direct effects of game-event–induced emotion shocks on emotional valence, the congruence effects are statistically significant in both cases of positive and negative game-event–induced emotion shocks. In sum, the findings presented in Table 2 still hold. We also match viewers in the treatment and control groups using propensity score matching (PSM) and obtain similar results.

Valid matching estimation rests on an unconfoundedness assumption that the treatment is independent of the outcome conditional on the matching variables, which is also known as the selection on observables. Put differently, unconfoundedness assumes that no unobserved variables exist that systematically affect whether the subjects receive a treatment and the outcome (Imbens 2004). Although we use CEM and propensity score matching to control for some observed demographic variables, they cannot rule out the possibility that some unobserved variables inherently make viewers selected to the treatment group more susceptible to television-program-induced emotion shocks. However, we cannot think of an unobserved variable that has such an effect.

6.3. Pretrend Analysis

A key identifying assumption of the DID analysis is that treatment and control groups have the same trend with respect to the outcome variable in the absence of treatment. Therefore, we collected a random sample of tweets that were posted by Broncos fans, Panthers fans, and fans of the other 30 NFL teams about one hour before the game to test whether the arousal and emotional valence levels were similar across *treated* and *control* groups.

Following Danaher et al. (2010), we also create relative time dummies to measure how far ahead of the game it was when the tweet was posted. Because the raw data are very noisy, we use a set of time dummies to indicate whether the tweet was posted earlier than 60, between 60 and 50, 50 and 40, 40 and 30, 30 and 20, 20 and 10, or 10 and 0 minutes ($\tau = -6, -5, -4, -3, -2, -1, 0$) before the game. We again estimate the following DID model:

$$\text{Log}(Volume)_{kt} = \sum_{\tau=-6}^{-1} \beta_\tau \mathbb{1}(P_t = \tau) + \sum_{\tau=-6}^{-1} \gamma_\tau \mathbb{1}(P_t = \tau) * Treat_k + \alpha_k + \epsilon_{kt}, \quad (3)$$

$$y_{ikt} = \sum_{\tau=-6}^{-1} \beta_\tau \mathbb{1}(P_t = \tau) + \sum_{\tau=-6}^{-1} \gamma_\tau \mathbb{1}(P_t = \tau) * Treat_i + \delta X_{it} + \alpha_k + \epsilon_{ikt}, \quad (4)$$

where $\text{Log}(Volume)_{kt}$ represents the logarithm of the per-minute tweet volume posted by a team k at time t,

Table 9. Fans of Other Teams as the Control Group with Positive Game Shocks

	Volume		Arousal		Valence	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	−0.804***	−0.851***	0.065***	−0.041	−0.016	0.033**
	(0.031)	(0.067)	(0.024)	(0.033)	(0.010)	(0.014)
Treat × *Post*	0.952***	0.884***	0.211***	0.133***	−0.009	−0.047
	(0.033)	(0.085)	(0.038)	(0.049)	(0.016)	(0.037)
Treat × *Congruent Ad Sentiment*		0.067		0.127***		−0.007
		(0.085)		(0.041)		(0.017)
Post × *Congruent Ad Sentiment*		0.054		0.131***		−0.060***
		(0.070)		(0.027)		(0.011)
Treat×*Post* × *Congruent Ad Sentiment*		0.079***		0.058***		0.044**
		(0.029)		(0.018)		(0.021)
Cumulative Game Sentiment	0.221***	0.220***	−0.085***	−0.082***	−0.064***	−0.057***
	(0.033)	(0.033)	(0.031)	(0.031)	(0.013)	(0.013)
Log(time after ad)	−0.397***	−0.397***	−0.036***	−0.028***	−0.031***	−0.034***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.004)	(0.004)
Log(followers)			−0.018*	−0.017*	0.003	0.003
			(0.009)	(0.009)	(0.004)	(0.004)
Log(followings)			−0.000	0.000	−0.012**	−0.012**
			(0.012)	(0.012)	(0.005)	(0.005)
Log(likes)			0.000	0.000	0.012***	0.011***
			(0.004)	(0.004)	(0.002)	(0.002)
Log(number of tweets)			−0.006	−0.007	−0.014***	−0.013***
			(0.007)	(0.007)	(0.003)	(0.003)
Ad fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted r^2	0.58	0.58	0.06	0.06	0.07	0.07
Observations	15,608	15,608	16,614	16,614	16,614	16,614

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The dependent variables are the logarithm of the volume of tweets posted by a certain group of users per minute, the arousal score of a tweet posted by a certain user, or the emotional valence of a tweet posted by a certain user, respectively. *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is not a fan of Broncos or Panthers, but is a fan of the other 30 NFL teams that did not make it to the Super Bowl. The *Post* takes the value of one if the tweet was posted after a positive game shock. The *Congruent Ad Sentiment* is an indicator which takes a value of one if the game-induced emotion is congruent with the ad-induced emotion, and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

and y_{ikt} represents the arousal or valence score of the tweet posted by user i at time t, who belongs to team k. $\boldsymbol{\alpha_k}$ is a vector of team fixed effects. The dummy variable $Treat_i$ equals one if the tweet was posted by Broncos' or Panthers' fans, and zero if it was posted by fans of any of the other 30 NFL teams. A vector of control variables $\boldsymbol{X_{it}}$ controls for some user characteristics such as a user's number of followers, followings, the total numbers of likes and tweets at time t as he/she posted the tweet. The dummy variable $\mathbb{1}(P_t = \tau)$ is a dummy variable which takes the value of 1 if it is period τ before the game started. For $\tau = -1$ to -6, it indicates the period between $(|\tau| + 1) * 10$ minutes and $|\tau| * 10$ minutes before the game. Therefore, $\mathbb{1}(\mathbf{P_t} = \boldsymbol{\tau})$ for $\tau = -1$ to $\tau = -6$ is a vector of time fixed effects. If the other 30 NFL teams are a good control group for the treated groups, we would expect $\gamma_\tau = 0$ for all periods before the game. Period $\tau = 0$ was used as the baseline period.

As shown in columns (1)–(3) in Table 12, no significant difference exists in the levels of arousal and valence between treatment and control groups in each pregame period, except the coefficients on *Treat* × *Period Minus 1* in columns (1) and (3) and the coefficient on *Treat* × *Period Minus 4* in column (2). We also conduct an F-test of the joint hypothesis that the true coefficients of all interaction terms, *Treat* × *Period Minus t*, equal zero in the model. The hypothesis cannot be rejected at the 1% level because the p values are 0.401, 0.362, and 0.682 for columns (1), (2), and (3), respectively. This lends further support to our common trend assumption. The positive coefficients on *Treat* × *Period Minus 1* in columns (1) and (3) might indicate that as time approached to the beginning of the game, the treatment group seemed to have higher valence in their tweets and posted slightly more compared with the control group.

6.4. Placebo Test

Our results in previous sections have shown consistent results across different model estimations. To lend

Table 10. Fans of Other Teams as the Control Group with Negative Game Shocks

	Volume		Arousal		Valence	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	−0.808***	−0.865***	0.070***	−0.069**	−0.015	0.036**
	(0.030)	(0.067)	(0.024)	(0.033)	(0.010)	(0.014)
Treat × *Post*	0.969***	0.910***	0.183***	−0.033	0.018	0.020
	(0.032)	(0.085)	(0.036)	(0.053)	(0.015)	(0.016)
Treat × *Congruent Ad Sentiment*		0.069		−0.276***		−0.007
		(0.085)		(0.068)		(0.029)
Post × *Congruent Ad Sentiment*		0.067		0.164***		−0.061***
		(0.069)		(0.028)		(0.012)
Treat×*Post* × *Congruent Ad Sentiment*		0.068**		0.235***		0.098**
		(0.033)		(0.050)		(0.043)
Cumulative Game Sentiment	−0.103***	−0.102***	0.105***	0.087***	0.052***	0.042***
	(0.033)	(0.033)	(0.030)	(0.031)	(0.013)	(0.013)
Log(time after ad)	−0.396***	−0.396***	−0.037***	−0.027***	−0.032***	−0.035***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.004)	(0.004)
Log(followers)			−0.027***	−0.025***	0.004	0.004
			(0.009)	(0.009)	(0.004)	(0.004)
Log(followings)			0.003	0.002	−0.016***	−0.016***
			(0.012)	(0.012)	(0.005)	(0.005)
Log(likes)			0.001	0.001	0.012***	0.012***
			(0.004)	(0.004)	(0.002)	(0.002)
Log(number of tweets)			−0.005	−0.006	−0.012***	−0.012***
			(0.007)	(0.007)	(0.003)	(0.003)
Ad fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted r^2	0.58	0.58	0.05	0.06	0.07	0.07
Observations	15,608	15,608	17,188	17,188	17,188	17,188

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The dependent variables are the logarithm of the volume of tweets posted by a certain group of users per minute, the arousal score of a tweet posted by a certain user, or the emotional valence of a tweet posted by a certain user, respectively. *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is not a fan of Broncos or Panthers, but is a fan of the other 30 NFL teams that did not make to the Super Bowl. The *Post* takes the value of one if the tweet was posted after a negative game shock. The *Congruent Ad Sentiment* is an indicator which takes a value of one if the game-induced emotion is congruent with the ad-induced emotion, and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

further support to the validity of our results, we also run a placebo test to check whether the game event–induced emotional shocks are picking up some spurious effects.

In our data set, we have 21 major game events that we identified to serve as emotion shocks. Our strategy is randomizing those 21 game events so that they are in a completely random order, to create a pseudo or placebo treatment. Using the pseudo treatment, we rerun our regression models in Equations (1) and (2). We replicate this procedure 100 times[17] and store the coefficient estimates. Repeating the shuffle provides a reliable check against outliers and helps ensure the robustness of the results. Table 13 presents the estimated average β_2 and β_5, along with their average standard deviations.

Results are presented in Table 13. Row (1) shows that the estimated average coefficient of $Treat_i \times Post_t$ is not significantly different from zero and therefore is the estimated average coefficient on the interaction term $Treat_i \times Post_t \times CongruentAdSentiment_j$ (row (3)). This result suggests that, with the randomly assigned treatment, there is no effect of television-program-induced emotion shocks on viewers' oWOM responses to ads. It lends further support to our earlier results and rules out the possibility that our estimates have picked up some spurious effects.

6.5. Alternative Valence Score

Our results are also robust to alternative measurement of the emotional valence, which is a continuous score between $[-1, 1]$, instead of a discrete binary measure used earlier. Results shown in Table 14 again indicate that negative game-induced emotion shocks resulted in a greater decline in the emotional valence

Table 11. Results on the Volume-Based and Text-Based Arousal and Valence After CEM

	Volume				Arousal				Valence			
	Positive game shock		Negative game shock		Positive game shock		Negative game shock		Positive game shock		Negative game shock	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post	–0.623***	–0.530***	–0.617***	–0.510***	0.028*	–0.050**	0.035**	–0.072***	–0.019***	0.036***	–0.016***	0.033***
	(0.031)	(0.068)	(0.031)	(0.068)	(0.015)	(0.020)	(0.015)	(0.020)	(0.006)	(0.008)	(0.006)	(0.008)
Treat × *Post*	0.539***	0.324***	0.556***	0.373***	0.147***	0.028	0.171***	0.003	–0.007	–0.059**	0.015	0.013
	(0.033)	(0.086)	(0.033)	(0.086)	(0.023)	(0.032)	(0.022)	(0.035)	(0.010)	(0.024)	(0.010)	(0.011)
Treat × *Congruent Ad Sentiment*		–0.331***		–0.328***		0.022		–0.246***		–0.014		–0.052***
		(0.086)		(0.086)		(0.024)		(0.044)		(0.011)		(0.018)
Post × *Congruent Ad Sentiment*		–0.108		–0.126*		0.095***		0.123***		–0.067***		–0.059***
		(0.070)		(0.070)		(0.016)		(0.017)		(0.007)		(0.007)
Treat × *Post*×*Congruent Ad Sentiment*		0.254***		0.216**		0.159***		0.212***		0.062**		0.128***
		(0.093)		(0.093)		(0.032)		(0.034)		(0.027)		(0.028)
Cumulative Game Sentiment	0.146***	0.141***	–0.120***	–0.115***	–0.053*	–0.056**	0.071***	0.052*	0.002	0.004	0.020***	0.016***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.028)	(0.028)	(0.027)	(0.027)	(0.005)	(0.005)	(0.005)	(0.005)
Log(time after ad)	–0.417***	–0.417***	–0.422***	–0.422***	–0.025***	–0.020***	–0.030***	–0.025***	–0.029***	–0.032***	–0.030***	–0.033***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Log(followers)					–0.010**	–0.009**	–0.015***	–0.015***	–0.001	–0.001	–0.002	–0.002
					(0.005)	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Log(followings)					–0.001	–0.002	0.003	0.004	0.001	0.000	0.001	0.001
					(0.005)	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Log(likes)					–0.001	–0.000	–0.001	–0.001	0.005***	0.005***	0.005***	0.005***
					(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Log(number of tweets)					–0.012***	–0.012***	–0.011***	–0.011***	–0.009***	–0.008***	–0.007***	–0.007***
					(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Ad fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted r^2	0.57	0.57	0.57	0.57	0.05	0.05	0.05	0.05	0.06	0.07	0.06	0.07
Observations	15,608	15,608	15,608	15,608	38,408	38,408	35,204	35,204	38,408	38,408	35,204	35,204

Notes. The results are based on a group of users who posted tweets about the ads shown during Super Bowl 2016. The dependent variables are the logarithm of the volume of tweets posted by a certain group of users, regarding a certain ad per minute, the arousal score of a tweet posted by a certain user, or the emotional valence of a tweet posted by a certain user, respectively. *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is neither a fan of Broncos or Panthers, nor of the other NFL teams. The *Post* takes the value of one if the tweet was posted after a positive or negative emotional shock. The *Congruent Ad Sentiment* is an indicator which takes a value of one if the game-event-induced emotional shock is congruent with the ad-induced emotion, and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

Table 12. Pretrend Analysis

	Volume	Arousal	Valence
	(1)	(2)	(3)
Treat × *Period Minus 1*	0.029*	0.025	0.105*
	(0.017)	(0.159)	(0.063)
Treat × *Period Minus 2*	–0.018	0.136	0.047
	(0.018)	(0.137)	(0.054)
Treat × *Period Minus 3*	–0.013	–0.101	–0.023
	(0.018)	(0.129)	(0.051)
Treat × *Period Minus 4*	0.019	0.191*	–0.003
	(0.018)	(0.115)	(0.046)
Treat × *Period Minus 5*	0.012	–0.040	–0.017
	(0.018)	(0.117)	(0.047)
Treat × *Period Minus 6*	–0.001	0.178	–0.011
	(0.018)	(0.121)	(0.048)
Period Minus 1	–0.073***	0.242***	0.037
	(0.014)	(0.071)	(0.028)
Period Minus 2	–0.008	0.066	–0.014
	(0.015)	(0.061)	(0.024)
Period Minus 3	–0.006	0.004	0.158***
	(0.015)	(0.064)	(0.025)
Period Minus 4	–0.032**	–0.043	0.120***
	(0.015)	(0.060)	(0.024)
Period Minus 5	–0.023	0.075	0.209***
	(0.015)	(0.062)	(0.025)
Period Minus 6	–0.004	0.077	0.134***
	(0.015)	(0.058)	(0.023)
Log(followers)		–0.028***	–0.001
		(0.006)	(0.002)
Log(followings)		0.028***	0.032***
		(0.007)	(0.003)
Log(likes)		0.002	–0.009***
		(0.003)	(0.001)
Log(number of tweets)		–0.009*	–0.010***
		(0.005)	(0.002)
Group fixed effects	Yes	Yes	Yes
Adjusted r^2	0.03	0.03	0.02
Observations	54,295	27,250	27,250

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The dependent variables are the logarithm of the volume of tweets posted by a certain group of users per minute, the arousal score of a tweet posted by a certain user, or the emotional valence of a tweet posted by a certain user, respectively. *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is a fan of any of the other 30 NFL teams that did not make to the Super Bowl. We include time fixed effects for the relative time dummies and interaction terms between *Treat* and the relative time dummies *Period Minus t*, which indicate the tweet was posted by the treated group between $t * 10$ and $(t-1) * 10$ minutes before the game. Group fixed effects are included. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

of the fan group's oWOM toward the ad, compared with the nonfan group, whereas positive game-induced emotion shocks seemed to have no significant effect on the emotional valence of viewers' oWOM toward the ad. When emotions induced by game events were congruent with ad-induced emotions, the fan group tended to have a more positive response to the ad compared with the incongruent scenario. This result is again consistent with our findings earlier.

We also performed a split-sample test by separating our sample into *Broncos' fans only* and *Panthers' fans only*, because systematic differences might exist between these two teams in their oWOM response to game-induced emotion shocks. The results presented in Tables A.4, A.5, and A.6 in Online Appendix D demonstrate similar findings.

6.6. Results on Super Bowl 2017

To examine whether our findings are only valid in this specific setting, we collected a whole new data set including viewers' tweets about ads, ads, and the game progress transcript for Super Bowl 2017 between New England Patriots and Atlanta Falcons. On February 5, 2017, New England Patriots defeated Atlanta Falcons, 34 to 28 in overtime. The Falcons led 20-9 in the late second quarter and 28-9 at the beginning of the final quarter. However, Patriots overcame the 25-point deficit, which was the largest comeback in the Super Bowl history and took the game into overtime. It was also the first Super Bowl that was decided in overtime. Nearly 112 million viewers watched the game, similar to Super Bowl 2016.[18]

Following a similar fashion, we used Twitter Streaming API to capture tweets posted during the game, from 6:30 p.m. to 10:45 p.m. Eastern Standard Time, on February 5, 2017. The keywords applied to the API are related to the 37 commercial advertisements aired during the game, including each ad's related hashtags.[19] For example, tweets that included #MoreInspired are considered to be related to Life Water's commercial. We first choose users who followed any of the other 30 NFL teams that did not make to the final as the control group. Then we followed the same CEM procedure to construct a matched sample for the treated group. The summary statistics and timeline are shown in Table A.3 and Figure A.3 in Online Appendix C.

Using a matched sample of users who followed any of the other 30 NFL teams as the control group, we repeated our DID analysis on each of these metrics of oWOM: arousal (measured by both volume and text) and valence. Results are displayed in Table 15. The significant and positive coefficients on $Treat_k \times Post_t$ in columns (1) and (3) again imply that game-induced emotion shocks significantly elevated the fan groups' arousal level than that of comparable control groups. However, coefficients on $Treat_k \times Post_t \times CongruentAdSentiment_j$ in columns (2) and (4) are not significant.

The coefficients on $Treat_i \times Post_t$ shown in columns (5) and (7) are again statistically significant and positive, consistent with our findings earlier in the

Table 13. Placebo Test

	Volume				Arousal				Valence			
Estimation	Positive		Negative		Positive		Negative		Positive		Negative	
Average *Treat* × *Post*	0.030	0.002	0.043	–0.002	–0.001	–0.011	–0.001	–0.010	0.001	0.006	0.001	0.005
Average standard error	0.027	0.059	0.027	0.061	0.012	0.019	0.013	0.020	0.005	0.008	0.005	0.008
Average *Treat* × *Post* × *CongruentAdSentiment*		0.034		0.054		0.029		0.031		–0.015		–0.014
Average standard error		0.062		0.065		0.030		0.032		0.012		0.013
Number of replications	100	100	100	100	100	100	100	100	100	100	100	100

case of Super Bowl 2016 (Hypothesis 1a is supported). Moreover, we only found partial support for our Hypothesis 2a, as indicated by the coefficient on $Treat_i \times Post_t \times CongruentAdSentiment_j$, which is statistically significant and positive in column (6), and not statistically significant in column (8). Therefore, Hypothesis 2a is only supported in the case of positive game-induced emotion shocks. Compared with previous results of Super Bowl 2016, one possible explanation is that the progress of Super Bowl 2017

Table 14. Results on Continuous Valence Score with Positive and Negative Emotion Shocks

	Positive emotional shock		Negative emotional shock	
	(1)	(2)	(3)	(4)
Post	0.009	0.060***	0.012	0.060***
	(0.007)	(0.009)	(0.007)	(0.010)
Treat × *Post*	–0.004	–0.050*	–0.026**	–0.035***
	(0.012)	(0.028)	(0.012)	(0.012)
Treat × *Congruent Ad Sentiment*		0.003		–0.009
		(0.013)		(0.022)
Post × *Congruent Ad Sentiment*		-0.063***		–0.059***
		(0.008)		(0.008)
Treat × *Post* × *Congruent Ad Sentiment*		0.048*		0.039**
		(0.027)		(0.017)
Cumulative Game Sentiment	–0.010	–0.008	0.022***	0.019***
	(0.006)	(0.006)	(0.006)	(0.006)
Log(time after ad)	0.002	–0.001	0.000	–0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Log(followers)	0.008***	0.007***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
Log(followings)	0.008***	0.008***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
Log(likes)	–0.009***	–0.009***	–0.009***	–0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Log(number of tweets)	–0.010***	–0.010***	–0.011***	–0.011***
	(0.002)	(0.002)	(0.001)	(0.001)
Ad fixed effects	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes
Adjusted r^2	0.09	0.09	0.09	0.09
Observations	41,930	41,930	43,072	43,072

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2016. The group of control users are those who do not support any NFL team. The dependent variable is the emotional valence of a tweet posted by a certain user, which is a continuous variable taking values between –1 and 1. The variable *Treat* takes the value of one for groups of users who are fans of either Broncos or Panthers, and zero if this user is not a fan of any NFL teams including Broncos or Panthers. *Post* takes the value of one if the tweet was posted after the exogenous emotional shock, such as a touchdown. The *Congruent Ad Sentiment* is an indicator which takes a value of one if the game-induced emotion is congruent with the ad-induced emotion, and zero if it is not. Standard errors are reported in parentheses. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

Table 15. Effects of Television-Program-Induced Emotion Shocks on the Volume-Based and Text-Based Arousal and Valence in Super Bowl 2017

	Volume				Arousal				Valence			
	Positive game shock		Negative game shock		Positive game shock		Negative game shock		Positive game shock		Negative game shock	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post	−0.624***	−0.698***	−0.836***	−0.999***	−0.092***	−0.133***	−0.090***	−0.131***	−0.021***	−0.017*	−0.020***	−0.017*
	(0.067)	(0.132)	(0.082)	(0.167)	(0.033)	(0.038)	(0.033)	(0.038)	(0.008)	(0.009)	(0.008)	(0.009)
Treat × *Post*	0.829***	0.809***	0.577***	0.651***	0.507***	−0.234	0.398***	0.665*	0.041	0.184*	0.013	0.043
	(0.070)	(0.150)	(0.086)	(0.192)	(0.150)	(0.452)	(0.152)	(0.382)	(0.035)	(0.105)	(0.035)	(0.088)
Treat × *Congruent Ad Sentiment*		0.195		0.310		−0.380		0.048		0.001		0.038
		(0.166)		(0.204)		(0.288)		(0.273)		(0.067)		(0.063)
Post × *Congruent Ad Sentiment*		0.102		0.212		0.070**		0.070**		−0.006		−0.006
		(0.145)		(0.182)		(0.033)		(0.033)		(0.008)		(0.008)
Treat × *Post* × *Congruent Ad Sentiment*		0.015		−0.107		0.813*		−0.343		−0.162		−0.031
		(0.170)		(0.215)		(0.478)		(0.421)		(0.111)		(0.097)
Cumulative Game Sentiment	−0.036***	−0.035***	0.022**	0.022**	−0.031	−0.025	−0.052	−0.045	0.006	0.006	0.002	0.002
	(0.005)	(0.005)	(0.010)	(0.010)	(0.037)	(0.037)	(0.035)	(0.035)	(0.008)	(0.009)	(0.008)	(0.008)
Log(time after ad)	−0.336***	−0.336***	−0.279***	−0.279***	0.005	0.006	0.003	0.004	−0.013***	−0.013***	−0.014***	−0.014***
	(0.012)	(0.012)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	(0.003)	(0.003)	(0.003)	(0.003)
Log(followers)					0.008	0.009	0.010	0.010	−0.002	−0.002	−0.002	−0.002
					(0.011)	(0.011)	(0.011)	(0.011)	(0.002)	(0.002)	(0.002)	(0.002)
Log(followings)					−0.005	−0.006	−0.005	−0.005	0.002	0.002	0.002	0.002
					(0.011)	(0.011)	(0.011)	(0.011)	(0.003)	(0.003)	(0.003)	(0.003)
Log(likes)					−0.004	−0.004	−0.003	−0.003	−0.001	−0.001	−0.001	−0.001
					(0.005)	(0.005)	(0.005)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)
Log(number of tweets)					−0.001	−0.002	−0.004	−0.003	0.001	0.001	0.001	0.001
					(0.006)	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)
Ad fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted r^2	0.68	0.68	0.74	0.74	0.25	0.25	0.25	0.25	0.04	0.04	0.04	0.04
Observations	25,245	25,245	6,594	6,594	8,678	8,678	8,519	8,519	8,678	8,678	8,519	8,519

Notes. The results are based on a group of users who posted tweets about ads shown during Super Bowl 2017. The dependent variable is volume-based arousal in terms of the logarithm of the volume of tweets posted by a certain group of users regarding a certain ad per minute, the arousal score of a tweet posted by a certain user, or the emotional valence of a tweet posted by a certain user, respectively. *Treat* takes the value of one for groups of users who are fans of either New England Patriots or Atlanta Falcons, and zero if this user is not a fan of Patriots or Falcons, but is a fan of the other 30 NFL teams. The *Post* takes the value of one if the tweet was posted after a major game event. The *Congruent Ad Sentiment* is an indicator that takes a value of one if the game-induced emotion is congruent with the ad-induced emotion, and zero if it is not. The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by group fixed effects and ad fixed effects. Standard errors are reported in parentheses.

***, **, and *Significance at the 1%, 5%, and 10% levels, respectively.

was quite different. In Super Bowl 2016, Denver Broncos outplayed Carolina Panthers from the very start of the game to the end of the game with a large comfortable lead. By contrast, Super Bowl 2017 was much breathtaking and dramatic, especially toward the last quarter of the game when the Patriots started to overcome the huge deficit and retake the game. Thus, for fans, their engagement with Super Bowl 2017 was much deeper and emotional. Thus, they might be less excited and aroused by ads because most of their attention was on the game according to the cognitive capacity theory (Mackie and Worth 1989, Lang et al. 1995), especially in the case of negative game shocks because of *negative* bias as we discussed earlier.

Furthermore, columns (9)–(12) present results on the effects of game-induced emotion shocks on the valence of tweets posted by users. The coefficients on $Treat_i \times Post_t$ are not statistically significant, and neither are the coefficients on $Treat_i \times Post_t \times CongruentAdSentiment_j$. Therefore, we did not find significant results. One possible explanation might be because of the small sample size that renders the significance of results small.

Overall, we find some results replicate the previous studies on Super Bowl 2016. The main difference is that Hypotheses 2a and 2b were strongly supported in Super Bowl 2016, whereas only partially supported (Hypothesis 2a) or not supported (Hypothesis 2b) in Super Bowl 2017.

6.7. Measurement Errors

Our dependent variables, namely arousal of tweets toward ads and valence of tweets toward ads, are constructed by algorithms. Although these algorithms tend to have high accuracy and can perform the corresponding tasks on large-scale data automatically, they are understandably imperfect. As a result, there could be potential measurement errors in these variables, which may be critical to model estimations. For example, if the classification error in a tweet's valence is correlated with whether the user is a fan of either team in the treatment group or of a team in the control group, it may cause biases when we estimate the effect of ad-induced and game-induced emotions. In this work, we take several approaches to alleviate this issue and ensure that the measurement error is not a major concern.

First, we use the lexicon-based method to measure the arousal of tweets. The lexicon is well established, and the arousal score is simply based on word frequency. Therefore, the only way measurement error can occur is when fans of the treated teams and control teams tend to use a different vocabulary to write their tweets that might result in systematically different arousal scores, or the lexicon is not comprehensive enough to include some arousal-related words in tweets that are disproportionally more frequently used by the treated or control team. To address this, a common way is to use alternative measures that are less prone to measurement errors. Following this, we use the volume-based approach in addition to the text-based approach to measure the arousal of tweets. We show that results using both measures are in general quite consistent as demonstrated in our empirical results.

Second, we use a supervised machine learning model (CNN-LSTM) to classify the valence of tweets. There are several factors that may affect the model's performance and cause measurement errors such as the model's learnability, high dimensionality in tweet data, and tweet's short length. Because we have the ground truth for valence (which was used to train and test our classifier), we can directly measure the measurement errors and check if they are correlated with the Twitter user's fanships. Specifically, we randomly sample 200 tweets per team from our data set. After that, we apply our CNN-LSTM classifier and measure the classification error. It is important to note that the classifier was previously trained and tested and we just apply it for out-of-sample prediction. Finally, we measure the Pearson correlation between the classification errors of tweets from fans of any two teams. This pair-wise correlation results are shown in Table 16. It is clear that the correlation is statistically insignificant for any pair of two teams regardless they are in treatment or control groups, indicating the noise introduced by measurement errors and team fanships are uncorrelated. We repeat the same procedure for examining the correlation between the classification errors of tweets from fans and nonfans. Similar to the results of the previous study, once again we do not find the correlation to be significant. Finally, we also generate four random samples (250 tweets each) from our data set. We measure pair-wise correlation between measurement errors of these samples and found similar results that the significance across samples are not significant (Table 17). Based on these results, the impact of measurement errors is of less concern to our studies.

7. General Discussion, Implications, and Suggestions for Further Research

As Lewis and Reiley (2013) urge future research to investigate how brand-related effects vary by audience, our research extends this line of research and examines how brand-related emotions vary by audience, even if the viewers watch the same ad on TV. Based on a multisource data set on social TV activities in Super Bowl 2016 and by using the exogenous nature of emotion-inducing game events in the Super Bowl, our results show causal links between the

Table 16. Correlations Between Classification Errors of Fans of NFL Teams

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1)	1.0																															
2)	−0.1	1.0																														
3)	0.3	−0.1	1.0																													
4)	−0.2	0.1	0.3	1.0																												
5)	−0.1	0.1	−0.3	0.2	1.0																											
6)	0.3	−0.2	−0.2	−0.2	0.1	1.0																										
7)	0.2	−0.2	0.0	0.1	−0.2	−0.3	1.0																									
8)	0.2	−0.2	−0.2	0.0	0.2	−0.1	−0.1	1.0																								
9)	0.0	−0.2	0.0	0.3	−0.1	0.0	0.3	−0.1	1.0																							
10)	−0.1	−0.2	−0.3	−0.1	−0.2	−0.3	−0.1	−0.2	−0.2	1.0																						
11)	0.3	0.0	0.3	−0.3	0.1	0.1	0.2	0.1	0.3	−0.1	1.0																					
12)	−0.2	0.0	0.1	0.0	0.1	−0.2	0.0	−0.2	0.1	−0.2	−0.1	1.0																				
13)	0.0	0.0	0.3	0.1	−0.2	0.3	0.1	−0.2	−0.2	−0.3	0.2	0.3	1.0																			
14)	−0.1	0.0	0.0	−0.3	0.0	−0.2	0.1	−0.1	0.2	0.0	−0.3	0.1	0.2	1.0																		
15)	−0.2	−0.1	0.2	0.3	0.2	−0.3	−0.2	0.2	0.1	−0.1	−0.2	0.3	−0.1	−0.2	1.0																	
16)	−0.2	0.0	0.0	−0.1	−0.2	0.2	0.1	0.2	−0.2	−0.1	0.3	−0.1	0.2	−0.1	0.2	1.0																
17)	−0.2	−0.2	0.2	0.1	0.0	0.0	0.2	−0.1	−0.3	−0.1	0.0	0.1	0.0	−0.1	−0.3	0.0	1.0															
18)	0.2	0.2	−0.2	−0.3	0.0	−0.3	0.2	0.3	−0.1	0.2	−0.1	0.1	−0.1	0.1	−0.3	0.2	0.2	1.0														
19)	0.0	0.2	0.3	0.2	−0.2	0.2	0.0	0.0	0.0	−0.2	−0.2	−0.1	0.0	0.1	0.1	0.2	0.1	0.3	1.0													
20)	−0.2	0.2	0.1	−0.2	0.1	−0.1	0.0	0.3	0.1	0.2	0.0	0.2	0.2	−0.3	0.0	−0.1	0.1	0.1	0.2	1.0												
21)	−0.1	−0.1	0.2	0.2	0.2	−0.1	0.2	−0.2	−0.1	−0.3	0.2	0.0	0.0	0.3	−0.3	−0.2	0.1	−0.2	0.1	0.2	1.0											
22)	0.1	−0.1	0.0	−0.1	0.1	0.3	−0.1	−0.3	−0.2	0.2	0.3	0.0	0.0	0.1	0.2	−0.3	−0.3	0.1	0.3	0.2	0.1	1.0										
23)	0.3	0.2	0.3	−0.1	−0.2	0.1	0.0	−0.2	−0.3	0.1	−0.2	0.3	0.1	0.3	−0.1	0.1	0.0	−0.1	−0.2	0.0	−0.1	−0.1	1.0									
24)	0.1	−0.3	0.1	0.2	0.3	0.0	0.2	0.2	−0.2	0.2	0.3	−0.1	0.1	−0.1	0.1	−0.1	0.1	0.0	0.0	−0.2	0.3	0.1	−0.1	1.0								
25)	−0.1	0.1	0.2	0.0	0.1	0.2	−0.1	0.2	0.2	0.1	0.3	−0.1	0.1	0.1	0.0	−0.1	−0.3	0.1	−0.2	0.1	0.3	−0.3	−0.2	−0.2	1.0							
26)	0.1	−0.1	0.2	0.1	0.0	0.2	0.1	−0.1	−0.2	−0.1	−0.1	−0.3	−0.3	0.3	0.0	0.2	0.2	−0.2	0.3	0.0	−0.3	−0.3	0.1	0.0	0.1	1.0						
27)	−0.3	−0.3	0.0	0.0	−0.3	0.3	−0.3	0.1	−0.2	0.3	0.2	−0.2	−0.2	−0.2	−0.2	−0.2	0.2	−0.1	−0.1	0.2	−0.2	−0.2	−0.2	0.1	0.3	−0.2	1.0					
28)	0.2	0.1	0.3	0.1	−0.1	−0.2	0.1	−0.3	0.0	−0.3	−0.2	0.3	−0.3	0.2	0.2	0.1	0.1	0.1	0.2	−0.2	−0.1	−0.2	0.3	−0.1	−0.2	0.0	0.0	1.0				
29)	−0.2	0.0	0.1	0.3	−0.1	−0.2	−0.2	0.3	0.2	0.0	−0.1	−0.2	−0.2	0.1	0.2	−0.2	0.2	−0.1	0.0	0.0	−0.1	−0.1	0.0	−0.3	0.1	−0.2	0.1	−0.2	1.0			
30)	0.2	0.1	0.0	0.2	0.2	−0.2	0.2	0.2	−0.1	−0.1	0.3	−0.1	0.3	−0.1	0.1	−0.2	−0.2	0.2	−0.1	−0.2	0.2	−0.1	0.0	0.0	0.3	−0.2	−0.2	−0.3	0.2	1.0		
31)	0.0	−0.2	0.2	−0.3	0.0	0.0	0.2	−0.3	−0.1	−0.1	−0.2	0.2	0.1	0.3	0.0	−0.2	−0.1	−0.3	−0.3	−0.1	−0.2	−0.1	0.0	−0.2	−0.3	0.1	−0.1	−0.2	−0.1	−0.2	1.0	
32)	0.1	−0.1	0.2	0.1	0.3	−0.1	−0.1	0.1	−0.1	−0.2	−0.1	−0.1	0.2	−0.3	0.1	0.2	−0.3	0.2	0.2	0.3	0.3	−0.1	0.3	−0.2	0.0	0.2	0.1	0.1	0.2	−0.2	0.1	1.0

Notes. We use Pearson correlation between classification errors of tweets from fans of any two teams, to check if measurement errors are correlated with Twitter user's fanships. None of these correlations is statistically significant (bold if significant) for any pair of two teams regardless they are in treatment or control groups, indicating the noise introduced by measurement errors and team fanships are uncorrelated. We repeat the same procedure for examining the correlation between the classification errors of tweets from fans and nonfans. The teams are (1) Arizona Cardinals, (2) Atlanta Falcons, (3) Baltimore Ravens, (4) Buffalo Bills, (5) Carolina Panthers, (6) Chicago Bears, (7) Cincinnati Bengals, (8) Cleveland Browns, (9) Dallas Cowboys, (10) Denver Broncos, (11) Detroit Lions, (12) Green Bay Packers, (13) Houston Texans, (14) Indianapolis Colts, (15) Jacksonville Jaguars, (16) Kansas City Chiefs, (17) Los Angeles Chargers, (18) Los Angeles Rams, (19) Miami Dolphins, (20) Minnesota Vikings, (21) New England Patriots, (22) New Orleans Saints, (23) New York Giants, (24) New York Jets, (25) Oakland Raiders, (26) Philadelphia Eagles, (27) Pittsburgh Steelers, (28) San Francisco 49ers, (29) Seattle Seahawks, (30) Tampa Bay Buccaneers, (31) Tennessee Titans, and (32) Washington Redskins.

Table 17. Correlations Between Classification Errors of Random Samples of Tweets from Fans of NFL Teams

	s1	s2	s3	s4
s1	1			
s2	0.173	1		
s3	0.259	0.035	1	
s4	−0.234	0.022	−0.102	1

Notes. We generate four random samples (250 tweets each) from our data set. We measure pair-wise Pearson correlation between the measurement errors of these samples and found that the correlations between classification errors across samples are not significant.

television-program-induced emotional context, ad-induced emotions, and viewers' oWOM activities toward ads. In our first set of results, we find that that the exogenous emotional context of the television program in which the ad airs can significantly influence the viewers' oWOM communication toward ads. More importantly, the match between television-program-induced emotions and ad-induced emotions can lead to a congruency effect, which intensifies the effect on the audience's arousal and valence in their oWOM toward ads.

Our findings confirm and extend prior research conducted in controlled laboratory experiments (Goldberg and Gorn 1987) to the context of social TV. In our study, we could find similar effects manifested in field settings as well, and individuals respond to advertising on social media platforms in a way that aligns with what theories predict. Moreover, our results also extend the recent work on social TV activity, which finds the connection between ads and the volume of WOM (Fossen and Schweidel 2016). Our finding is also consistent with previous reports that argue ads with negative emotions may lead to an increment in the volume of oWOM about ads (Agrawal 2016, Berger 2016). Because of the heightened arousal level, we find that both positive and negative emotions lead to an increase in the volume of oWOM about ads. In addition, our results suggest that the interaction between ad and television program may also influence the arousal and emotional valence of online WOM.

Our results suggest a number of media planning strategies and ad design strategies that advertisers, marketers, and television networks can implement to influence the online brand WOM. Our results also have implications for oWOM behavior. We discuss these implications of our key results.

7.1. Theoretical Implications

The theoretical contributions of this research are multifold. First, this study provides empirical evidence on how television-program-induced emotions can both directly and jointly affect viewers' social TV activities along with ad-induced emotions. To date, prior literature on social TV and cross-media effects has predominantly focused on the relationship between oWOM and TV advertising only, but has yet to investigate the interplay between television programs and advertising, despite the fact that ads are embedded in TV programs (Hill et al. 2012, Fossen and Schweidel 2016). To the best of our knowledge, this work is among the first in the literature to fill the gap by exploring the joint relationship between oWOM, TV advertising, and TV programs. Our study implicated that the ad, when properly aligned with TV program's emotional context, can provide a mood-lifting opportunity to significantly improve the audience's ad perception and evaluation, as reflected in the audience's mood arousal states and valence. Thus, these results shed light on how advertisers, marketers, and television networks can use program characteristics and ad characteristics to boost oWOM for their respective ads. We further illustrate these in the following section.

Second, compared with research on the effects on viewers' valence in social TV (Diakopoulos and Shamma 2010, Zhao et al. 2011, Yu and Wang 2015), the effects on viewers' arousal remain scant, let alone on viewers' arousal toward TV advertising. We contribute to the literature by providing evidence indicating a positive relationship between television program–invoked emotions and viewers' arousal in their oWOM toward the ad. Importantly, this effect may interact with emotional cues from ads, such that when emotional cues invoked from the television program are matched with those from ads, the effect on viewers' arousal toward ads may be enhanced. This study contributes to prior research that studies effects of mood on ad perception, expanding the scope of examination beyond valence.

Third, our work also contributes to the literature on oWOM (King et al. 2014). A notable research thread in oWOM is to explore factors that motivate and affect consumers to engage in WOM. Thus far, most well-studied factors are so-called individual factors such as self-enhancement (Fiske 2002, Wojnicki and Godes 2008, De Angelis et al. 2012), innovativeness and opinion leadership (Sun et al. 2006), self-efficacy (Gruen et al. 2006), individuation (Ho and Dempsey 2010), and altruism (Hennig-Thurau et al. 2003, Dellarocas and Narayan 2006). In other words, these factors are about user's own characteristics and thus endogenous to the WOM behavior. However, little research has examined exogenous factors as potential drivers, partly because of lack of information about users who generate oWOM. The only exception we can find is the work by Bakhshi et al. (2014), which found weather (e.g., temperature, rain, snow, season) can affect online review ratings and polarity. On the other hand, previous studies found situational factors can significantly affect traditional WOM. For example, Lau and Ng (2001) identified the proximity between two individuals could positively affect negative WOM, which reduces the likelihood of product purchase. Although WOM and oWOM are two different types of communication (WOM is conveyed face-to-face, whereas online WOM is computer mediated), the participants of the communication (regardless of WOM and oWOM) are inevitably exposed to and affected by the physical and emotional environment they are in. Thus, this work tries to fill this important research gap in oWOM by examining a largely unexplored yet ubiquitous situational factor: a users' emotional states and their role in users' oWOM communications. Our results showed that users' emotional states indeed have a significant influence on the arousal and valence of their online WOM communications.

Finally, this paper contributes to the literature on advertising strategies by proposing a new method based on social TV activities to examine viewers' reactions to ads. Compared with traditional methods used in previous studies such as surveys, interviews, and laboratory experiments (Batra and Stayman 1990, Brader 2006), our method provides an immediate measure of the impact of television advertising, which is both instantaneous and more cost-efficient.

7.2. Pratical Implications

Our work provides some clear and practical implications, as it suggests that advertisers and television networks should be cognizant of the emotional states of advertisements and programming content when placing ads. Our results suggest how networks distribute ads in programs can have a meaningful impact on the arousal and valence of oWOM about ads. Currently, advertisers on media buying strategies mostly focus on the order of ad positions in program breaks (Katz 2016). Our results here provide a different perspective to the current practice by showing that it is also important to consider the context of the program in which the ad airs. Following this, networks may want to optimally allocate ads into different programs or different positions in a certain program, depending on the ad content and the emotional context induced by the program. The networks shall be aware that viewers who are exposed to a negative emotional context can have a significantly more negative reaction toward ads. However, a congruency effect may help mitigate such effects. It is worth noting that our findings might not help inform live TV advertising, as in the Super Bowl, but in prerecorded TV, where planning ahead is more feasible, aligning ad emotion with TV content emotion could render some synergies. In sum, networks can use the findings in this paper to offer more program-specific recommendations to advertisers who are interested in finding the right balance among the television program, the messages that the ad wants to deliver, and online brand chatter.

Our study also has implications for ad copy testing, which determines an ad's effectiveness based on consumer responses, feedback, and behavior (Alwitt and Prabhaker 1992). Testing an ad in a particular context or without any context can be a potentially biased procedure. It may lead to misleading results, depending upon the type of consumers tested, the congruency between the ad and the context, and the appreciation of the latter. Therefore, one implication of this study is that ads should be tested with specific target groups in mind in terms of involvement and context appreciation. This is becoming important given people are increasingly expressing their thoughts and views on social media platforms about ads, products and so on. They also heavily rely on such oWOM to make their purchase choices.

In addition to advertising on TV, our findings also have implications for general ad strategies on other mediums and platforms. For instance, most Internet ads are placed through a complex auction process in which an advertiser bid for better placements of its ads to be clicked on (Edelman et al. 2007). The early online placement did not consider the context of the ad. Recently, a new area called contextual advertising has become increasingly popular and important in both industry and academia (Schmitter and Rosen 2005, Broder et al. 2007, Henkin et al. 2017). Contextual advertising is a form of targeted advertising for advertisements appearing on websites or other media. The advertisements themselves are selected and served by automated systems based on the context of what a user is looking at.[20] One example is Google targets ads based on page content.[21] Previous studies have shown the importance of context appreciation for advertising effectiveness (De Pelsmacker et al. 2002). Thus, in order to have an effective placement of online ads, there is a need for embedding ads in highly appreciated contexts. Our findings support the suggestion made by Lynch and Stipp (1999) that advertisers should take situational factors such as viewers' emotional states into account when designing the ad placement/bidding system. Thus, the direct implication is that the context of ads and the content of ads should be more carefully matched or contrasted, especially when ads are likely to stimulate positive feedbacks from the audience (e.g., ads on sport equipment on a web page about a local baseball team won the playoff spot). Meanwhile, for ad platforms such as Google and Facebook, they may also consider the potential effect of matching between contexts and ads and incorporate it as a part of ad their auction strategies to maximize profits.

7.3. Final Remarks

First and foremost, having an understanding of how television-program-induced emotions affect oWOM toward ads is fundamentally important for advertisers and marketers to maximize the effectiveness of television advertising campaigns. We show that social TV can be valuable as it can result in a massive amount of oWOM for advertised brands. This chatter creates free exposures for the brand online, extends the reach of television ad campaigns to the online space, and offers real-time feedback to advertisers to examine the effectiveness of their ads. Given the prevalence of the social TV activities, we believe it is important to have a better understanding of the interplay between television-program-induced emotions and advertising-induced emotions. Our research also demonstrates that advertisers could not only use social media as a marketing tool to promote their brands and engage their customers but also extract insightful consumer information from social media. Our results suggest a number of media planning strategies related to the emotional context that television networks can implement to impact online brand WOM.

Although the results of this study shed some light on the effect of television-program-induced emotions on online brand chatter, our study has several limitations that may weaken our ability to precisely quantify the magnitude of the effect. First, tweets are often short and noisy. In order to measure a viewer's sentiment more precisely, we tried several methods and achieved an accuracy of 91.5% (see Online Appendix B). Future studies can further focus on improving the accuracy in sentiment measurement. Second, although we have a control group that consists of viewers who do not root for either team (or any NFL team), judged based on whether they are following these two teams (or any NFL team) on Twitter, it might still be possible that the viewers take sides with one team during the game watching. If that is the case, although they are not avid football fans, they may be subject to the influence of program-induced emotions. However, we think our robustness checks including using followers of the other 30 NFL teams could mitigate this concern to some extent, because it is generally uncommon for a user to follow two NFL teams at the same time. Third, self-selection into WOM transmission is a common problem for most research that uses oWOM data, because oWOM is essentially a self-reporting behavior. As suggested by Zhu and Zhang (2006), Dellarocas et al. (2007), and Li and Hitt (2008), oWOM is subject to the self-selection bias, similar to the survey method. Unless we adopt a modeling approach to specifically correct the self-selection bias, it remains challenging to fully resolve this issue.[22] Besides, there might exist other casual mechanisms that produce the effects we observe. For instance, one user's WOM may indirectly influence another users's emotional response in oWOM. However, we could not find a way to tease out this possibility given the data limitation. Last, the Super Bowl is a specific game context, and its viewers may be of specific characteristics. Therefore, assessing the generalizability of our findings to other contexts would be of interest in future work.

Endnotes

[1] See http://www.statista.com/statistics/216526/super-bowl-us-tv-viewership/.

[2] We thank an anonymous reviewer for suggesting this paper.

[3] See http://www.espn.com/nfl/game?gameId=400820438.

[4] See https://en.wikipedia.org/wiki/Television_timeout.

[5] In order to avoid the possibility of multiple emotional shocks and eliminate other confounding factors that might influence viewers' response to the ad, we limit viewers' tweets to those that were posted within five minutes after the ad was shown. We also made sure between each of ads-related tweets and its associated ad's airing, at most one game event took place.

[6] See http://www.espn.com/nfl/playbyplay?gameId=400820438.

[7] See https://ftw.usatoday.com/2017/02/super-bowl-espn-win-probability-atlanta-falcons-new-england-patriots-stats-tom-brady.

[8] Another example is Mountain Dew's Super Bowl ad that features a bizarre hybrid character called "Puppy, Monkey, Baby." After its debut, this ad became a trending topic on Twitter and received many critics. More controversial ads, such as the humorous Doritos' Super Bowl ad "Ultrasound," may inspire disgust, whereas other people may find them uproariously funny. See https://en.wikipedia.org/wiki/Puppy_monkey_baby.

[9] In the subsequent analysis using other 30 NFL teams, k is the group indicator for each of these 30 NFL teams and these two teams in Super Bowl.

[10] We also tried a 15-, 10-, and 3-minute window. The results are similar.

[11] In our analysis, we separately analyze the effects of positive game-induced emotion shocks and the negative game-induced emotion shocks.

[12] The main effects of *Treat* and *Congruent Ad Sentiment* have been subsumed by the group fixed effects and ad fixed effects.

[13] A within-user analysis is not feasible because of the few number of tweets that a user posts.

[14] Other possible explanations might be that the *positive* event (e.g., a touchdown) might even have an emotional effect on users who are not fans of either team. We thank the associate editor for suggesting this idea.

[15] We thank an anonymous reviewer and the associate editor for suggesting this idea.

[16] We thank an anonymous reviewer for this great suggestion.

[17] See http://www.espn.com/nfl/game?gameId=400927752.

[18] If no hashtags exist, we use a brand's Twitter official account and the brand's name.

[19] See https://en.wikipedia.org/wiki/Contextual_advertising.

[20] See https://support.google.com/adwords/answer/2497832?hl=en.

[21] On a related note, our findings might not generalize to the population that do not use Twitter. However, given our focus on examining effects of program-induced emotions on viewers' oWOM behavior, we think it is interesting to investigate effects of program-induced emotions on the offline WOM in future studies.

References

Adomavicius G, Tuzhilin A (2015) *Context-Aware Recommender Systems. Recommender Systems Handbook* (Springer, New York).

Aggarwal VA, Hsu DH (2013) Entrepreneurial exits and innovation. *Management Sci.* 60(4):867–887.

Agrawal A (2016) The pros and cons of controversial marketing. Accessed, https://www.forbes.com/sites/ajagrawal/2016/01/08/the-pros-and-cons-of-controversial-marketing/#3ecb1fbb3e51/.

Alwitt LF, Prabhaker PR (1992) Functional and belief dimensions of attitudes to television advertising: Implications for copytesting. *J. Advertising. Res.* 32(5):30–42.

Averill JR (1969) Autonomic response patterns during sadness and mirth. *Psychophysiology* 5(4):399–414.

Babin BJ, Darden WR, Babin LA (1998) Negative emotions in marketing research: Affect or artifact? *J. Bus. Res.* 42(3):271–285.

Bakhshi S, Kanuparthy P, Gilbert E (2014) Demographics, weather and online reviews: A study of restaurant recommendations. *Proc. 23rd Internat. Conf. on World Wide Web* (ACM, New York), 443–454.

Batra R, Stayman DM (1990) The role of mood in advertising effectiveness. *J. Consumer Res.* 17(2):203–214.

Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev. General Psych.* 5(4):323.

Benton A, Hill S (2012) The Spoiler Effect?: Designing social tv content that promotes ongoing WOM. *INFORMS Conf. Information Systems and Technology* (Citeseer).

Berger J (2011) Arousal increases social transmission of information. *Psych. Sci.* 22(7):891–893.

Berger J (2016) *Contagious: Why Things Catch On* (Simon and Schuster, New York).

Berger J, Sorensen AT, Rasmussen SJ (2010) Positive effects of negative publicity: When negative reviews increase sales. *Marketing Sci.* 29(5):815–827.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1): 249–275.

Botha E, Reyneke M (2013) To share or not to share: the role of content and emotion in viral marketing. *J. Public Affairs* 13(2): 160–171.

Brader T (2006) *Campaigning for Hearts and Minds: How Emotional Appeals in Political Ads Work* (University of Chicago Press, Chicago).

Bradley MM, Lang PJ (1999) Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, Center for Research in Psychophysiology, University of Florida. 30(1):25–36.

Broder A, Fontoura M, Josifovski V, Riedel L (2007) A semantic approach to contextual advertising. *Proc. 30th Annual Internat. ACM SIGIR Conf. Research Development Inform. Retrieval* (ACM, New York), 559–566.

Cantor JR, Zillmann D, Bryant J (1975) Enhancement of experienced sexual arousal in response to erotic stimuli through misattribution of unrelated residual excitation. *J. Personality Soc. Psych.* 32(1):69.

Carlson M, Miller N (1987) Explanation of the relation between negative mood and helping. *Psych. Bull.* 102(1):91.

Chang Y (2019) Spectators' emotional responses in tweets during the super bowl 50 game. *Sport Management Rev.* 22(3):348–362.

Chen Y, Xie J (2008) Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Sci.* 54(3):477–491.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.

Cialdini RB, Kenrick DT, Baumann DJ (1982) Effects of mood on prosocial behavior in children and adults. *Development of Prosocial Behavior* (Academic Press), 339–359.

Danaher B, Dhanasobhon S, Smith MD, Telang R (2010) Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and Internet piracy. *Marketing Sci.* 29(6):1138–1151.

De Angelis M, Bonezzi A, Peluso AM, Rucker DD, Costabile M (2012) On braggarts and gossips: A self-enhancement account of word-of-mouth generation and transmission. *J. Marketing Res.* 49(4): 551–563.

De Pelsmacker P, Geuens M, Anckaert P (2002) Media context and advertising effectiveness: The role of context appreciation and context/ad similarity. *J. Advertising* 31(2):49–61.

Dellarocas C, Narayan R (2006) What motivates consumers to review a product online? a study of the product-specific antecedents of online movie reviews. *Workshop on Information Systems and Economics (WISE)*, 1–6.

Dellarocas C, Zhang X, Awad NF (2007) Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *J. Interactive Marketing* 21(4):23–45.

Diakopoulos NA, Shamma DA (2010) Characterizing debate performance via aggregated twitter sentiment. *Proc. SIGCHI Conf. Human Factors Comput. Systems*, (ACM, New York) 1195–1198.

Ding X, Liu B, Yu PS (2008) A holistic lexicon-based approach to opinion mining. *Proc. 2008 Internat. Conf. Web Search Data Mining* (ACM New York), 231–240.

Duan W, Gu B, Whinston AB (2008) Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems* 45(4):1007–1016.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Elder RS, Krishna A (2011) The "visual depiction effect" in advertising: Facilitating embodied mental simulation through product orientation. *J. Consumer Res.* 38(6):988–1003.

Eroglu SA, Machleit KA, Davis LM (2001) Atmospheric qualities of online retailing: A conceptual model and implications. *J. Bus. Res.* 54(2):177–184.

Fiedler K, Nickel S, Muehlfriedel T, Unkelbach C (2001) Is mood congruency an effect of genuine memory or response bias? *J. Experiment. Soc. Psych.* 37(3):201–214.

Fiske AP (2002) Using individualism and collectivism to compare cultures–a critique of the validity and measurement of the constructs: Comment on Oyserman et al. (2002). *Psych. Bull.* 128(1):78–88.

Forgas JP, Bower GH (1987) Mood effects on person-perception judgments. *J. Personality Soc. Psych.* 53(1):53.

Forman C, Ghose A, Goldfarb A (2009) Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Sci.* 55(1):47–57.

Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inform. Systems Res.* 19(3):291–313.

Fossen BL, Schweidel DA (2016) Television advertising and online word-of-mouth: An empirical investigation of social TV activity. *Marketing Sci.* 36(1):105–123.

Frederick EL, Lim CH, Clavio G, Walsh P (2012) Why we follow: An examination of parasocial interaction and fan motivations for following athlete archetypes on twitter. *Internat. J. Sport Community* 5(4):481–502.

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Marketing Sci.* 23(4):545–560.

Goldberg ME, Gorn GJ (1987) Happy and sad TV programs: How they affect reactions to commercials. *J. Consumer Res.* 14(3): 387–403.

Gonzalez G, De La Rosa JL, Montaner M, Delfin S (2007) Embedding emotional context in recommender systems. *Proc. 23rd Internat. Conf. Data Engrg. Workshop* (IEEE, New York), 845–852.

Gray J, Sandvoss C, Harrington CL (2017) *Fandom: Identities and Communities in a Mediated World* (NYU Press, New York).

Gruen TW, Osmonbekov T, Czaplewski AJ (2006) ewom: The impact of customer-to-customer online know-how exchange on customer value and loyalty. *J. Bus. Res.* 59(4):449–456.

Ha Y, Im H (2012) Role of website design quality in satisfaction and word of mouth generation. *J. Service Management* 23(1):79–96.

Henkin A, Shaham Y, Brickner I (2017) Contextual advertising techniques for implemented at mobile devices. US Patent 9,710,818.

Hennig-Thurau T, Walsh G, Walsh G (2003) Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the Internet. *Internat. J. Electronic Commerce* 8(2): 51–74.

Hill S, Nalavade A, Benton A (2012) Social TV: Real-time social media response to TV advertising. *Proc. 6th Internat. Workshop Data Mining Online Advertising Internet Econom.* (ACM, New York), 4.

Hinojosa JA, Carretié L, Méndez-Bértolo C, Míguez A, Pozo MA (2009) Arousal contributions to affective priming: Electrophysiological correlates. *Emotion* 9(2):164.

Hinz O, Hill S, Kim JY, et al (2016) TV's dirty little secret: The negative effect of popular TV on online auction sales. *MIS Quart* 40(3): 623–644.

Ho JY, Dempsey M (2010) Viral marketing: Motivations to forward online content. *J. Bus. Res.* 63(9-10):1000–1006.

Holbrook MB, Batra R (1987) Assessing the role of emotions as mediators of consumer responses to advertising. *J. Consumer Res.* 14(3):404–420.

Iacus SM, King G, Porro G (2011) Causal inference without balance checking: Coarsened exact matching. *Political Anal.* 20(1):1–24.

Imbens GW (2004) Nonparametric estimation of average treatment effects under exogeneity: A review. *Rev. Econom. Statist.* 86(1): 4–29.

Isen AM, Shalker TE, Clark M, Karp L (1978) Affect, accessibility of material in memory, and behavior: A cognitive loop? *J. Personality Soc. Psych.* 36(1):1.

Kamins MA, Marks LJ, Skinner D (1991) Television commercial evaluation in the context of program induced mood: Congruency vs. consistency effects. *J. Advertising* 20(2):1–14.

Katz HE (2016) *The Media Handbook: A Complete Guide to Advertising Media Selection, Planning, Research, and Buying* (Taylor & Francis, Boca Raton, FL).

Kim AJ, Johnson KK (2016) Power of consumers using social media: Examining the influences of brand-related user-generated content on facebook. *Comput. Human Behav.* 58:98–108.

King RA, Racherla P, Bush VD (2014) What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *J. Interactive Marketing* 28(3):167–183.

Kuperman V, Estes Z, Brysbaert M, Warriner AB (2014) Emotion and language: Valence and arousal affect word recognition. *J. Experiment. Psych. General* 143(3):1065.

Kwark Y, Chen J, Raghunathan S (2014) Online product reviews: Implications for retailers and competing manufacturers. *Inform. Systems Res.* 25(1):93–110.

Ladhari R (2007) The effect of consumption emotions on satisfaction and word-of-mouth communications. *Psych. Marketing* 24(12): 1085–1108.

Lang A, Dhillon K, Dong Q (1995) The effects of emotional arousal and valence on television viewers' cognitive capacity and memory. *J. Broadcasting Electronic Media* 39(3):313–327.

Lau GT, Ng S (2001) Individual and situational factors influencing negative word-of-mouth behaviour. *Canadian J. Admin. Sci.* 18(3):163–178.

Lewis RA, Reiley DH (2013) Down-to-the-minute effects of super bowl advertising on online search behavior. *Proc. 14th ACM Conf. Electronic Commerce* (ACM, New York), 639–656.

Lewis RA, Reiley DH (2014) Online ads and offline sales: Measuring the effect of retail advertising via a controlled experiment on yahoo! *Quantitative Marketing Econom.* 12(3):235–266.

Li X, Hitt LM (2008) Self-selection and information role of online product reviews. *Inform. Systems Res.* 19(4):456–474.

Lynch K, Stipp H (1999) Examination of qualitative viewing factors for optimal advertising strategies. *J. Advertising Res.* 39(3):7.

Mackie DM, Worth LT (1989) Processing deficits and the mediation of positive affect in persuasion. *J. Personality Soc. Psych.* 57(1):27.

Mehrabian A, Russell JA (1974) An Approach to Environmental Psychology (MIT Press, Cambridge, MA).

Midha A (2014) Study: Exposure to TV tweets drives consumers to take action: Both on and off of Twitter. https://blog.twitter.com/2014/study-exposure-to-tv-tweets-drives-consumers-to-take-action-both-on-and-off-of-twitter.

Mohammad SM (2016) *Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. Emotion Measurement* (Elsevier, New York).

Mudambi SM, Schuff D (2010) Research note: What makes a helpful online review? a study of customer reviews on amazon.com. *MIS Quart.* 34(1):185–200.

Myers DG (2004) Theories of emotion. *Psychology*, 7th ed. (Worth Publishers, New York), 500.

Nielsen (2015) Social TV: A bellwether for tv audience engagement. https://www.nielsen.com/us/en/insights/news/2015/social-tv-a-bellwether-for-tv-audience-engagement.html/.

Norris CE, Colman AM (1992) Context effects on recall and recognition of magazine advertisements. *J. Advertising* 21(3): 37–46.

Pavelchak MA, Antil JH, Munch JM (1988) The super bowl: An investigation into the relationship among program context, emotional experience, and ad recall. *J. Consumer Res.* 15(3):360–367.

Peng C, Kim YG (2014) Application of the stimuli-organism-response (sor) framework to online shopping behavior. *J. Internet Commerce* 13(3-4):159–176.

Perry SD, Jenzowsky SA, King CM, Yi H, Hester JB, Gartenschlaeger J (1997) Using humorous programs as a vehicle for humorous commercials. *J. Comm.* 47(1):20–39.

Pugh SD (2001) Service with a smile: Emotional contagion in the service encounter. *Acad. Management J.* 44(5):1018–1027.

Rajadesingan A, Zafarani R, Liu H (2015) Sarcasm detection on Twitter: A behavioral modeling approach. *Proc. 8th ACM Internat. Conf. Web Search Data Mining* (ACM, New York), 97–106.

Rozin P, Royzman EB (2001) Negativity bias, negativity dominance, and contagion. *Personality Soc. Psych. Rev.* 5(4):296–320.

Russell JA (1980) A circumplex model of affect. *J. Personality Soc. Psych.* 39(6):1161.

Schmitter T, Rosen J (2005) Contextual advertising system. U.S. Patent Application 10/836,820.

Singh SN, Churchill GA Jr (1987) Arousal and advertising effectiveness. *J. Advertising* 16(1):4–40.

Stevenson RA, Mikels JA, James TW (2007) Characterization of the affective norms for English words by discrete emotional categories. *Behav. Res. Methods* 39(4):1020–1024.

Sun T, Youn S, Wu G, Kuntaraporn M (2006) Online word-of-mouth (or mouse): An exploration of its antecedents and consequences. *J. Comput. Mediated Comm.* 11(4):1104–1127.

Tan C, Niculae V, Danescu-Niculescu-Mizil C, Lee L (2016) Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. *Proc. 25th Internat. Conf. World Wide Web* (ACM, New York), 613–624.

Tavassoli N, Shultz CJ, Fitzsimons GJ (1995) Program involvement: Are moderate levels best for ad memory and attitude toward the ad? *J. Advertising Res.* 35:61–72.

Toubia O, Stephen AT (2013) Intrinsic vs. image-related utility in social media: Why do people contribute content to Twitter? *Marketing Sci.* 32(3):368–392.

Tsang MM, Ho SC, Liang TP (2004) Consumer attitudes toward mobile advertising: An empirical study. *Internat. J. Electronic Commerce* 8(3):65–78.

Tyler BD, Cobbs J (2017) All rivals are not equal: Clarifying misrepresentations and discerning three core properties of rivalry. *J. Sport Management* 31(1):1–14.

Van Setten M, Pokraev S, Koolwaaij J (2004) Context-aware recommendations in the mobile tourist application compass. *Proc. Internat. Conf. Adaptive Hypermedia Adaptive Web-Based Systems* (Springer, New York), 235–244.

Wang C, Zhang P, Choi R, D'Eredita M (2002) Understanding consumers attitude toward advertising. *Proc. 8th Amer. Conf. Inform. Systems* (AIS, Atlanta), 1143–1148.

Warriner AB, Kuperman V, Brysbaert M (2013) Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* 45(4):1191–1207.

Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Personality Soc. Psych.* 54(6):1063.

Wojnicki AC, Godes D (2008) Word-of-mouth as self-enhancement. HBS Marketing Research Paper 06-01, http://ssrn.com/abstract=908999.

Yang X, Smith RE (2009) Beyond attention effects: Modeling the persuasive and emotional effects of advertising creativity. *Marketing Sci.* 28(5):935–949.

Yu Y, Wang X (2015) World Cup 2014 in the Twitter world: A big data analysis of sentiments in us sports fans' tweets. *Comput. Human Behav.* 48:392–400.

Zentner A, Smith M, Kaya C (2013) How video rental patterns change as consumers move online. *Management Sci.* 59(11): 2622–2634.

Zhao S, Zhong L, Wickramasuriya J, Vasudevan V (2011) Analyzing Twitter for social TV: Sentiment extraction for sports. *Proc. 2nd Internat. Workshop Future Television*, vol. 2, 11–18.

Zhu F, Zhang X (2006) The influence of online consumer reviews on the demand for experience goods: The case of video games. *ICIS 2006 Proc.* (AIS, Atlanta), 25.